

=> fil reg ; d que 13

FILE REGISTRY ENTERED AT 08:54:17 ON 06 JUL 2001
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2001 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 5 JUL 2001 HIGHEST RN 344737-88-8
DICTIONARY FILE UPDATES: 5 JUL 2001 HIGHEST RN 344737-88-8

TSCA INFORMATION NOW CURRENT THROUGH January 11, 2001

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

Structure search limits have been increased. See HELP SLIMIT
for details.

L3 31 SEA FILE=REGISTRY ABB=ON ^.{0-6}HK[ALIVPFWMSTYNQCG]K.{0-6}^/SQ
SP

=> d rn cn sql kwic nte 13 1-31; fil capl; d que nos 14

L3 ANSWER 1 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 344422-23-7 REGISTRY - use Registry # to match sequence to citation
CN L-Lysine, L-histidyl-L-lysyl-L-seryl-L-lysyl- (9CI) (CA INDEX NAME)
SQL 5

SEQ 1 HKSKK-lys
=====

HITS AT: 1-5

0-12
0,1

L3 ANSWER 2 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 338970-56-2 REGISTRY
CN L-Lysine, glycyl-L-histidyl-L-lysyl-L-alanyl-L-lysylglycyl-L-prolyl-L-
arginyl- (9CI) (CA INDEX NAME)
SQL 9

SEQ 1 GHKAKGPRK
=====

HITS AT: 1-9

L3 ANSWER 3 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 338970-52-8 REGISTRY
CN Glycine, glycyl-L-histidyl-L-lysyl-L-valyl-L-lysyl-L-arginyl-L-prolyl-L-
lysyl- (9CI) (CA INDEX NAME)
SQL 9

SEQ 1 GHKVKRPKG
=====

HITS AT: 1-9

L3 ANSWER 4 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 334993-69-0 REGISTRY
CN L-Lysine, L-.alpha.-glutamyl-L-histidyl-L-lysyl-L-valyl-L-lysyl-L-
isoleucylglycyl-L-valyl-L-.alpha.-glutamyl-L-glutaminyl- (9CI) (CA INDEX
NAME)
SQL 11

SEQ 1 EHKVKIGVEQ K

HITS AT: 1-11

L3 ANSWER 5 OF 31 REGISTRY COPYRIGHT 2001 ACS
 RN 316821-90-6 REGISTRY
 CN L-Lysine, L-lysyl-L-histidyl-L-lysyl-L-histidyl-L-lysyl-L-histidyl-L-lysylglycyl-L-lysyl-L-histidyl-L-lysyl-L-histidyl-L-lysyl-L-histidyl-
 (9CI) (CA INDEX NAME)

SQL 15

SEQ 1 KHKHKHKGKH KHKHK

HITS AT: 1-15

L3 ANSWER 6 OF 31 REGISTRY COPYRIGHT 2001 ACS
 RN 290374-35-5 REGISTRY
 CN L-Lysine, N2-acetyl-L-lysyl-L-threonyl-L-.alpha.-glutamyl-L-seryl-L-histidyl-L-lysyl-L-alanyl-L-lysylglycyl- (9CI) (CA INDEX NAME)

SQL 10

SEQ 1 KTESHKAKGK

HITS AT: 1-10

NTE modified

type	location	description
terminal mod.	Lys-1	N-acetyl

L3 ANSWER 7 OF 31 REGISTRY COPYRIGHT 2001 ACS
 RN 282712-87-2 REGISTRY
 CN L-Histidinamide, L-alanyl-L-lysyl-L-arginyl-L-histidyl-L-histidyl-L-lysyl-L-tyrosyl-L-lysyl-L-arginyl-L-lysyl-L-phenylalanyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 22: PN: W00040965 PAGE: 11 claimed protein

SQL 12

SEQ 1 AKRHHKYKRK FH

HITS AT: 1-12

NTE modified

type	location	description
terminal mod.	His-12	C-terminal amide

L3 ANSWER 8 OF 31 REGISTRY COPYRIGHT 2001 ACS
 RN 278169-06-5 REGISTRY
 CN L-Serine, L-arginyl-L-histidyl-L-lysyl-L-glutaminyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-alanyl-L-prolyl- (9CI) (CA INDEX NAME)

SQL 10

SEQ 1 RHKQKIIAPS

HITS AT: 1-10

L3 ANSWER 9 OF 31 REGISTRY COPYRIGHT 2001 ACS
 RN 251651-17-9 REGISTRY
 CN L-Valine, L-methionyl-L-seryl-L-arginyl-L-lysyl-L-histidyl-L-lysyl-L-

tryptophyl-L-lysyl-L-leucyl-L-serylglycyl- (9CI) (CA INDEX NAME)
SQL 12

SEQ 1 MSRKHKWKLS GV
=====

HITS AT: 1-12

L3 ANSWER 10 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 225518-96-7 REGISTRY
CN L-Threonine, L-histidyl-L-lysylglycyl-L-lysyl-L-lysyl-L-.alpha.-aspartyl-L-threonyl-L-seryl-L-isoleucyl- (9CI) (CA INDEX NAME)
SQL 10

SEQ 1 HKGKKDTSIT
=====

HITS AT: 1-10

L3 ANSWER 11 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 201799-95-3 REGISTRY
CN L-Leucine, L-.alpha.-aspartyl-L-lysyl-L-histidyl-L-lysyl-L-leucyl-L-lysyl-L-lysyl-L-seryl-L-.alpha.-glutamyl- (9CI) (CA INDEX NAME)
SQL 10

SEQ 1 DKHKLKSEL
=====

HITS AT: 1-10

L3 ANSWER 12 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 201799-94-2 REGISTRY
CN L-Lysine, L-arginyl-L-.alpha.-glutamylglycyl-L-.alpha.-aspartyl-L-lysyl-L-histidyl-L-lysyl-L-leucyl-L-lysyl- (9CI) (CA INDEX NAME)
SQL 10

SEQ 1 REGDKHKLKK
=====

HITS AT: 1-10

L3 ANSWER 13 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 187816-25-7 REGISTRY
CN L-Alanine, L-.alpha.-glutamyl-L-alanyl-L-prolyl-L-histidyl-L-lysyl-L-phenylalanyl-L-lysyl-L-seryl- (9CI) (CA INDEX NAME)
SQL 9

SEQ * 1 EAPHKFKSA
=====

HITS AT: 1-9

L3 ANSWER 14 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 187816-07-5 REGISTRY
CN L-Valine, L-.alpha.-glutamyl-L-alanyl-L-prolyl-L-histidyl-L-lysyl-L-phenylalanyl-L-lysyl-L-asparaginy- (9CI) (CA INDEX NAME)
SQL 9

SEQ 1 EAPHKFKNV
=====

HITS AT: 1-9

L3 ANSWER 15 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 186258-89-9 REGISTRY
CN L-Histidine, L-alanyl-L-lysyl-L-arginyl-L-histidyl-L-histidyl-L-lysyl-L-tyrosyl-L-lysyl-L-arginyl-L-lysyl-L-phenylalanyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 15: PN: WO0040965 PAGE: 19 claimed protein

SQL 12

SEQ 1 AKRHHKYKRK FH

=====

HITS AT: 1-12

L3 ANSWER 16 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 176260-60-9 REGISTRY

CN L-Lysine, N2-[N-[1-[N-[N-[N2-[N2-[N2-[N-(N2-L-cysteinyl-L-arginyl)-L-histidyl]-L-lysyl]-L-glutamyl]-L-lysyl]-L-isoleucyl]-L-isoleucyl]-L-alanyl]-L-prolyl]-L-alanyl]- (9CI) (CA INDEX NAME)

SQL 12

SEQ 1 CRHKQKIIAP AK

=====

HITS AT: 1-12

L3 ANSWER 17 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 176260-59-6 REGISTRY

CN L-Alanine, N-[1-[N-[N-[N2-[N2-[N2-[N-(N-L-cysteinyl-L-cysteinyl)-L-arginyl]-L-histidyl]-L-lysyl]-L-glutamyl]-L-lysyl]-L-isoleucyl]-L-isoleucyl]-L-alanyl]-L-prolyl]- (9CI) (CA INDEX NAME)

SQL 12

SEQ 1 CCRHKQKIIA PA

=====

HITS AT: 1-12

L3 ANSWER 18 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 161174-90-9 REGISTRY

CN L-Tyrosinamide, L-cysteinyl-L-histidyl-L-lysyl-L-leucyl-L-lysyl-L-alanyl-L-alanyl-L-leucyl-L-cysteinyl-, cyclic (1.fwdarw.9)-disulfide (9CI) (CA INDEX NAME)

SQL 10

SEQ 1 CHKLKAALCY

=====

HITS AT: 1-10

NTE modified

type	location	description
terminal mod.	Tyr-10	- C-terminal amide
bridge	Cys-1	- Cys-9 disulfide bridge

L3 ANSWER 19 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 161174-87-4 REGISTRY

CN L-Tyrosinamide, L-cysteinyl-L-histidyl-L-lysyl-L-leucyl-L-lysyl-L-alanyl-L-alanyl-L-leucyl-L-cysteinyl- (9CI) (CA INDEX NAME)

SQL 10

SEQ 1 CHKLKAALCY

=====

HITS AT: 1-10

NTE modified

type	location	description
------	----------	-------------

terminal mod. Tyr-10 - C-terminal amide

L3 ANSWER 20 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 157566-35-3 REGISTRY

CN L-Glutamic acid, L-arginyl-L-.alpha.-glutamylglycyl-L-.alpha.-aspartyl-L-lysyl-L-histidyl-L-lysyl-L-leucyl-L-lysyl-L-lysyl-L-seryl-L-.alpha.-glutamyl-L-leucyl-L-lysyl- (9CI) (CA INDEX NAME)

SQL 15

SEQ 1 REGDKHKLKK SELKE

=====

HITS AT: 1-15

L3 ANSWER 21 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 144776-43-2 REGISTRY

CN L-Lysine, N2-[N-[N2-[N-[N-(N2-L-arginyl-L-arginyl)-L-leucyl]-L-histidyl]-L-lysyl]-L-leucyl]- (9CI) (CA INDEX NAME)

SQL 7

SEQ 1 RRLHKLK

=====

HITS AT: 1-7

L3 ANSWER 22 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 144776-41-0 REGISTRY

CN L-Lysine, N2-[N-[N2-[N-[N-(N2-L-arginyl-L-arginyl)-L-.alpha.-glutamyl]-L-histidyl]-L-lysyl]-L-leucyl]- (9CI) (CA INDEX NAME)

SQL 7

SEQ 1 RREHKLK

=====

HITS AT: 1-7

L3 ANSWER 23 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 144776-40-9 REGISTRY

CN L-Lysine, N2-[N-[N2-[N-[N-(N2-L-arginyl-L-arginyl)-L-seryl]-L-histidyl]-L-lysyl]-L-leucyl]- (9CI) (CA INDEX NAME)

SQL 7

SEQ 1 RRS HKLK

=====

HITS AT: 1-7

L3 ANSWER 24 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 121795-39-9 REGISTRY

CN L-Cysteine, L-prolyl-L-threonyl-L-.alpha.-glutamyl-L-alanyl-L-arginyl-L-histidyl-L-lysyl-L-glutaminyl-L-lysyl-L-isoleucyl-L-valyl-L-alanyl-L-prolyl-L-valyl- (9CI) (CA INDEX NAME)

SQL 15

SEQ 1 PTEARHKQKI VAPVC

=====

HITS AT: 1-15

L3 ANSWER 25 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 115154-42-2 REGISTRY

CN L-Lysine, N2-[N-[N2-[N-[N2-[N-[N-(N-L-.alpha.-aspartyl-L-seryl)-L-histidyl]-L-histidyl]-L-lysyl]-L-alanyl]-L-lysyl]-L-alanyl]- (9CI) (CA INDEX NAME)

SQL 9

SEQ 1 DSHHKAKAK

=====

HITS AT: 1-9

L3 ANSWER 26 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 115154-41-1 REGISTRY

CN L-Lysine, L-seryl-L-histidyl-L-histidyl-L-lysyl-L-alanyl-L-lysylglycyl-
(9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN L-Lysine, N2-[N-[N2-[N-[N2-[N-(N-L-seryl-L-histidyl)-L-histidyl]-L-lysyl]-
L-alanyl]-L-lysyl]glycyl]-

SQL 8

SEQ 1 SHHKAKGK

=====

HITS AT: 1-8

L3 ANSWER 27 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 112424-93-8 REGISTRY

CN D-Alaninamide, N-acetyl-3-(2-naphthalenyl)-D-alanyl-4-chloro-D-
phenylalanyl-3-(3-pyridinyl)-D-alanyl-L-seryl-L-histidyl-N6-(3-
pyridinylcarbonyl)-D-lysyl-L-leucyl-N6-(3-pyridinylcarbonyl)-L-lysyl-L-
prolyl- (9CI) (CA INDEX NAME)

SQL 10

SEQ 1 AFASHKLKPA

=====

HITS AT: 1-10

NTE modified

type	location		description
terminal mod.	Ala-1	-	N-acetyl
terminal mod.	Ala-10	-	C-terminal amide
modification	Ala-1	-	2-naphthalenyl<2-Naph>
modification	Phe-2	-	chloro<Cl>
modification	Ala-3	-	3-pyridinyl<3Py>
modification	Lys-6	-	3-pyridinylcarbonyl
modification	Lys-8	-	3-pyridinylcarbonyl

L3 ANSWER 28 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 97461-89-7 REGISTRY

CN L-Lysine, N2-[1-(N2-L-histidyl-L-lysyl)-L-prolyl]- (9CI) (CA INDEX NAME)

SQL 4

SEQ 1 HKPK

=====

HITS AT: 1-4

L3 ANSWER 29 OF 31 REGISTRY COPYRIGHT 2001 ACS

RN 92269-40-4 REGISTRY

CN L-Isoleucine, N-[N2-[N2-[N2-(N-L-arginyl-L-histidyl)-L-lysyl]-L-
glutaminyl]-L-lysyl]- (9CI) (CA INDEX NAME)

SQL 6

SEQ 1 RHKQKI

=====

HITS AT: 1-6

L3 ANSWER 30 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 92269-39-1 REGISTRY
CN L-Lysine, N2-[N2-[N2-[N-(N2-L-alanyl-L-arginyl)-L-histidyl]-L-lysyl]-L-glutaminy]- (9CI) (CA INDEX NAME)
SQL 6

SEQ 1 ARHKQK
=====

HITS AT: 1-6

L3 ANSWER 31 OF 31 REGISTRY COPYRIGHT 2001 ACS
RN 92227-36-6 REGISTRY
CN L-Valine, N-[N-[N2-[N2-(N2-L-histidyl-L-lysyl)-L-glutaminy]-L-lysyl]-L-isoleucyl]- (9CI) (CA INDEX NAME)
SQL 6

SEQ 1 HKQKIV
=====

HITS AT: 1-6

FILE 'CAPLUS' ENTERED AT 08:54:46 ON 06 JUL 2001
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2001 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications.

FILE COVERS 1947 - 6 Jul 2001 VOL 135 ISS 3
FILE LAST UPDATED: 5 Jul 2001 (20010705/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

This file supports REGISTRY for direct browsing and searching of all substance data from the REGISTRY file. Enter HELP FIRST for more information.

CAPLUS now provides online access to patents and literature covered in CA from 1947 to the present. On April 22, 2001, bibliographic information and abstracts were added for over 2.2 million references published in CA from 1947 to 1966.

The CA Lexicon is now available in the Controlled Term (/CT) field. Enter HELP LEXICON for full details.

Attention, the CA Lexicon is the copyrighted intellectual property of the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

L3 31 SEA FILE=REGISTRY ABB=ON ^.{0-6}HK[ALIVPFWMSTYNQCG]K.{0-6}^/SQ
SP

L4 25 SEA FILE=CAPLUS ABB=ON L3 *Reg³try file answer set crossed over into CAPLUS to get citations*

=> d ibib abs hitrn 14 1-25; fil hom

L4 ANSWER 1 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2001:429545 CAPLUS
 TITLE: Peptides and peptidomimetics with structural similarity to human p53 that activate p53 function
 INVENTOR(S): Halazonetis, Thanos; Hartwig, Wolfgang
 PATENT ASSIGNEE(S): Bayer Corp., USA; The Wistar Institute
 SOURCE: U.S., 22 pp., Cont.-in-part of U.S. 6,169,073.
 CODEN: USXXAM
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 2
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6245886	B1	20010612	US 1997-894327	19971204
US 6169073	B1	20010102	US 1995-392542	19950216
WO 9625434	A1	19960822	WO 1996-US1535	19960216

W: AU, CA, JP, KR, US
 RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE

PRIORITY APPLN. INFO.: US 1995-392542 A2 19950216
 WO 1996-US1535 W 19960216

AB The present invention provides peptides and peptidomimetics corresponding to part or to the entirety of the region encompassed by residues 360-386 of human p53, said peptides and peptidomimetics characterized by the ability to activate DNA binding of wild-type p53 and to select tumor-derived p53 mutants. Pharmaceutical compns. of the compds. of the invention and methods of using these compns. therapeutically are also provided.

IT 344422-23-7 - *Use Registry # to match citation to sequence*
 RL: PRP (Properties)
 (unclaimed sequence; peptides and peptidomimetics with structural similarity to human p53 that activate p53 function)

REFERENCE COUNT: 34
 REFERENCE(S): (1) Anon; EP 0518650 1992 CAPLUS
 (2) Anon; WO 94/08241 1994 CAPLUS
 (3) Anon; WO 94/10306 1994 CAPLUS
 (4) Anon; WO 94/12202 1994 CAPLUS
 (5) Anon; WO 95/17213 1995 CAPLUS
 ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 2 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2001:284116 CAPLUS
 DOCUMENT NUMBER: 134:306975
 TITLE: Recombinant factor C and factor C fragments for endotoxin detection or removal and for use as anti-microbials
 INVENTOR(S): Ding, Jeak Ling; Ho, Bow; Tan, Nguan Soon
 PATENT ASSIGNEE(S): National University of Singapore, Singapore
 SOURCE: PCT Int. Appl., 123 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

 WO 2001027289 A2 20010419 WO 2000-SG162 20001013
 W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
 CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
 HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT,
 LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU,
 SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN,
 YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
 RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY,
 DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ,
 CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.:

US 1999-159569 P 19991015

US 2000-626795 A 20000726

AB Recombinant fragments of Factor C are disclosed. These proteins and peptides show great potency in recognizing, binding to, neutralizing, and removing endotoxin. These mols. can thus be used for anti-microbial, anti-endotoxin, and anti-sepsis therapy. SSCrFCES is a 38 kDa protein representing the LPS-binding domain of Factor C. The ability of SSCrFCES to bind lipid A was analyzed using an ELISA-based assay as well as surface plasmon resonance. Surface plasmon resonance similarly carried out for SSCrFC-sushi-1,2,3-GFP, SSCrFC-sushi-1-GFP, and SSCrFC-sushi-3-GFP confirmed their superior affinity for endotoxin. The 50 % endotoxin-neutralizing concn. of SSCrFCES against 200 EU of endotoxin is 0.069 μ M, suggesting that SSCrFCES is an effective inhibitor of Limulus amoebocyte lysis coagulation cascade. Although partially attenuated by human serum, as low as 1 μ M of SSCrFCES inhibits the LPS-induced secretion of hTNF- α and hIL-8 THP-1 and human peripheral blood mononuclear cells with a potency more superior than polymyxin B. SSCrFCES is non-cytotoxic, with a clearance rate of 4.7 mL/min. The LD90 of SSCrFCES for LPS lethality in mice is achieved at 2 μ M. These results demonstrate the endotoxin-neutralizing capability of SSCrFCES in vitro and in vivo, as well as its potential for use in the treatment of endotoxin-induced septic shock. Also embodied in this application is the use of the sushi peptides and their mutant derivs. as potent antimicrobials. Further embodied in this application is the use of sushi peptides or sushi recombinant proteins to remove endotoxin from liqs.

IT 334993-69-0P

RL: ARG (Analytical reagent use); BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); PREP (Preparation); USES (Uses)

(Sushi 3 peptide; recombinant factor C and factor C fragments for endotoxin detection or removal and for use as anti-microbials)

L4 ANSWER 3 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2001:211598 CAPLUS

DOCUMENT NUMBER: 134:351765

TITLE: Recombinant human adenovirus: targeting to the human transferrin receptor improves gene transfer to brain microcapillary endothelium

AUTHOR(S): Xia, Haibin; Anderson, Brian; Mao, Qinwen; Davidson, Beverly L.

CORPORATE SOURCE: Program in Gene Therapy, Departments of Internal Medicine, University of Iowa College of Medicine, Iowa City, IA, 52242, USA

SOURCE: J. Virol. (2000), 74(23), 11359-11366

CODEN: JOVIAM; ISSN: 0022-538X

PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Some inborn errors of metab. due to deficiencies of sol. lysosomal enzymes

cause global neurodegenerative disease. Representative examples include the infantile and late infantile forms of the ceroid lipofuscinoses (CLN1 or CLN2 deficiency, resp.) and mucopolysaccharidoses type VII (MPS VII), a deficiency of .beta.-glucuronidase. Treatment of the central nervous system component of these disorders will require widespread protein or enzyme replacement, either through dissemination of the protein or through dissemination of a gene encoding it. We hypothesize that transduction of brain microcapillary endothelium (BME) with recombinant viral vectors, with secretion of enzyme product basolaterally, could allow for widespread enzyme dissemination. To achieve this, viruses should be modified to target the BME. This requires (i) identification of a BME-resident target receptor, (ii) identification of motifs targeted to that mol., (iii) the construction of modified viruses to allow for binding to the target receptor, and (iv) demonstrated transduction of receptor-expressing cells. In proof of principal expts., we chose the human transferrin receptor (hTfR), a mol. found at high d. on human BME. A nonamer phage display library was panned for motifs which could bind hTfR. Forty-three clones were sequenced, most of which contained an AKxxK/R, KxKxPK/R, or KxK motif. Ten peptides representative of the three motifs were cloned into the HI loop of adenovirus type 5 fiber. All motifs tested retained their ability to trimerize and bind transferrin receptor, and seven allowed for recombinant adenovirus prodn. Importantly, the fiber-modified viruses facilitated increased gene transfer (2- to 34-fold) to hTfR expressing cell lines and human brain microcapillary endothelia expressing high levels of endogenous receptor. Our data indicate that adenoviruses can be modified in the HI loop for expanded tropism to the hTfR.

IT 338970-52-8P 338970-56-2P

RL: BPR (Biological process); SPN (Synthetic preparation); BIOL (Biological study); PREP (Preparation); PROC (Process).
(recombinant human adenovirus: targeting to the human transferrin receptor improves gene transfer to brain microcapillary endothelium)

REFERENCE COUNT: 36

REFERENCE(S): (1) Anderson, R; Gene Ther 2000, V7, P1034 CAPLUS
(2) Bosch, A; Hum Gene Ther 2000, V11, P1139 CAPLUS
(4) Broadwell, R; Exp Neurol 1996, V142, P47 CAPLUS
(6) Chartier, C; J Virol 1996, V70, P4805 CAPLUS
(8) Douglas, J; Nat Biotechnol 1996, V14, P1574 CAPLUS
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 4 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2000:728698 CAPLUS

DOCUMENT NUMBER: 134:91012

TITLE: Co-polymer of histidine and lysine markedly enhances transfection efficiency of liposomes

AUTHOR(S): Chen, Q-R.; Zhang, L.; Stass, S. A.; Mixson, A. J.

CORPORATE SOURCE: Department of Pathology and Greenebaum Cancer Center, University of Maryland Baltimore, Baltimore, MD, 21201, USA

SOURCE: Gene Ther. (2000), 7(19), 1698-1705

CODEN: GETHEC; ISSN: 0969-7128

PUBLISHER: Nature Publishing Group

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Development of nonviral delivery systems is progressing toward a transfection efficiency sufficient to affect metabolic and neoplastic diseases in humans. Nevertheless, inadequate transfection efficiency of target cells with current nonviral systems still limits the utility of this therapy. In the current study, we have detd. that a co-polymer of histidine and lysine (H-K) enhances the transfection efficiency of liposomes, a leading nonviral system. We found that in the absence of serum, the addn. of this polymer increased transfection as much as 10-fold

in comparison with the liposome:DNA complex alone. More impressively, the co-polymer in the presence of serum increased transfection efficiency up to 100-fold. Furthermore, in vivo expression of luciferase in a tumor increased 15-fold with the addn. of H-K polymer to the liposome:plasmid DNA complexes. Without liposomes, the H-K polymer had little to no effect on transfection efficiency. We anticipate that further modifications of this co-polymer will yield mols. with both increased complexity and transfection efficiency.

IT **316821-90-6P**

RL: BPR (Biological process); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses)

(co-polymer of histidine and lysine enhances transfection efficiency of liposomes)

REFERENCE COUNT: 35

REFERENCE(S): (1) Behr, J; Proc Natl Acad Sci USA 1989, V86, P6982 CAPLUS
(2) Boussif, O; Proc Natl Acad Sci USA 1995, V92, P7297 CAPLUS
(3) Budker, V; Nat Biotechnol 1996, V14, P760 CAPLUS
(4) Cayot, P; Anal Biochem 1997, V249, P184 CAPLUS
(5) Chen, Q; Cancer Res 1999, V59, P3308 CAPLUS
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 5 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2000:475506 CAPLUS

DOCUMENT NUMBER: 133:115150

TITLE: Methods using a histatin, histatin fragment, or related peptide for treating cystic fibrosis

INVENTOR(S): Spacciapoli, Peter; Rothstein, David M.; Friden, Phillip M.

PATENT ASSIGNEE(S): Periodontix, Inc., USA

SOURCE: PCT Int. Appl., 27 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000040204	A2	20000713	WO 2000-US480	20000107
W: AU, CA, NZ				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				

PRIORITY APPLN. INFO.: US 1999-226666 A 19990108

OTHER SOURCE(S): MARPAT 133:115150

AB Methods are disclosed for treating cystic fibrosis in a mammal that include administering an effective amt. of a histatin, a histatin fragment, or a histatin-related peptide.

IT **186258-89-9 282712-87-2**

RL: BAC (Biological activity or effector, except adverse); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses) (histatin, histatin fragment, or related peptide for treating cystic fibrosis)

L4 ANSWER 6 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2000:386830 CAPLUS

DOCUMENT NUMBER: 133:218720

TITLE: Ni(II) Specifically Cleaves the C-Terminal Tail of the Major Variant of Histone H2A and Forms an Oxidative

Damage-Mediating Complex with the Cleaved-Off Octapeptide

AUTHOR(S): Bal, Wojciech; Liang, Rongti; Lukszo, Jan; Lee, Sang-Han; Dizdaroglu, Miral; Kasprzak, Kazimierz S.

CORPORATE SOURCE: Laboratory of Comparative Carcinogenesis, National Cancer Institute, Frederick, MD, 21702, USA

SOURCE: Chem. Res. Toxicol. (2000), 13(7), 616-624
CODEN: CRTOEC; ISSN: 0893-228X

PUBLISHER: American Chemical Society

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The acetyl-TESHHK-amide peptide, modeling a part of the C-terminal "tail" of histone H2A, was found previously by us to undergo at pH 7.4 a Ni(II)-assisted hydrolysis of the E-S peptide bond with formation of a stronger Ni(II) complex with the SHHK-amide product (1998). To further characterize the hydrolysis and test the resulting Ni(II) complex for redox activity, bovine histone H2A and three peptides were investigated: acetyl-LLGKVTIAQGGVLPNIQAVLLPKKTESHHKAKGK (H2A34), modeling the entire "C-tail" of H2A; SHHKAKGK (H2A8), modeling the cutoff product of hydrolysis; and acetyl-KTESHHKAKGK (H2A10), modeling a putative Ni(II) binding site in a minor variant H2A.4 of human histone H2A. The Ni(II)-assisted hydrolysis of H2A and H2A34 was found to proceed approx. 7-fold faster than that of the Ni(II)-acetyl-TESHHK-amide complex under comparable conditions. In both cases, the Ni(II) complex with H2A8 was the smaller product of the hydrolysis, indicating a high site specificity of the reaction. Of three other metals tested with H2A34, only Cu(II) cleaved the E-S bond, although much less efficiently than Ni(II); Co(II) and Zn(II) had no effect whatsoever. The H2A10 peptide appeared to be fully resistant to hydrolytic cleavage and did not exhibit any redox activity vs. H2O2 in the presence of Ni(II) at pH 7.4. Likewise, redox-inactive was the Ni(II)-H2A34 complex. In contrast, the Ni(II)-H2A8 complex promoted oxidative damage of pUC19 DNA by H2O2, evidenced by a significant increase in the no. of single strand breaks and nucleobase modifications typical for a hydroxyl radical-like species attack on DNA. Interestingly, instead of 8-oxopurines, the corresponding formamidopyrimidines were the major products of the damage. The difference in redox activity between the Ni(II)-H2A34 and Ni(II)-H2A8 complexes is most likely assocd. with their different geometries: octahedral and square planar, resp. Incubation of the Ni(II)-H2A8 complex with H2O2 also resulted in degrdn. of the peptide ligand, esp. at its Ser and His residues. Thus, binding of Ni(II) to the ESHHK motif of the histone H2A C-tail is damaging to the histone C-terminal tail and to histone-assocd. DNA. The results support a dual mechanism of Ni(II)-induced carcinogenesis, including both genotoxic and epigenetic effects.

IT 115154-41-1 290374-35-5

RL: BSU (Biological study, unclassified); RCT (Reactant); BIOL (Biological study)

(Ni(II) specifically cleaves C-terminal tail of major variant of histone H2A and forms oxidative damage-mediating complex with cleaved-off octapeptide)

REFERENCE COUNT:

54

REFERENCE(S):

- (1) Arents, G; Proc Natl Acad Sci USA 1991, V88, P10148 CAPLUS
 - (2) Arents, G; Proc Natl Acad Sci USA 1995, V92, P11170 CAPLUS
 - (3) Bal, W; Arch Biochem Biophys 1999, V364, P161 CAPLUS
 - (4) Bal, W; Chem Res Toxicol 1995, V8, P683 CAPLUS
 - (5) Bal, W; Chem Res Toxicol 1997, V10, P915 CAPLUS
- ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 7 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2000:295355 CAPLUS
DOCUMENT NUMBER: 133:103515
TITLE: Analysis of foreign epitope inserts in HBcAg.
approaches to solving the problem of core particle
self-assembly
AUTHOR(S): Karpenko, L. I.; Ivanisenko, V. A.; Pika, I. S.;
Chikaev, N. A.; Eroshkin, A. M.; Melamed, N. V.;
Veremeiko, T. A.; Il'ichev, A. A.
CORPORATE SOURCE: Institute of Biotechnology, State Research Center
VECTOR, Novosibirsk, 633159, Russia
SOURCE: Mol. Biol. (2000), 34(2), 194-199
CODEN: MOLBBJ; ISSN: 0026-8933
PUBLISHER: MAIK Nauka/Interperiodica Publishing
DOCUMENT TYPE: Journal.
LANGUAGE: English

AB The hepatitis B virus core antigen (HBcAg) is a promising protein carrier
for exposing the epitopes of various human and animal pathogens.
HBcAg-based chimeric proteins can be used in creating highly efficient
vaccines; however, not all chimeric HBcAg with foreign epitope inserts are
capable of assembly into virus-like particles. Using computer programs
ProAnalyst, SALIX, and QSARPro, the authors examd. the relation between
the self-assembly capability of chimeric HBcAg and the physicochem.
properties of the inserts. The self-assembly was impaired when the
inserted peptides contained highly hydrophobic and bulky residues tending
to form .beta.-structures; this esp. concerned the C-proximal residues in
the insert. Recommendations were elaborated for constructing foreign
epitopes that would ensure correct self-assembly of chimeric HBcAg
particles.

IT 278169-06-5D, fusion products with hepatitis B core antigen
RL: BPR (Biological process); PRP (Properties); BIOL (Biological study);
PROC (Process)
(self-assembly core antigen particles expressing foreign epitopes in
hepatitis B core antigen)

REFERENCE COUNT: 26

REFERENCE(S): (1) Argos, P; EMBO J 1988, V7, P819 CAPLUS
(2) Bogardt, R; J Mol Evol 1980, V15, P197 CAPLUS
(3) Borisova, G; J Virology 1993, V67, P3696 CAPLUS
(4) Boulter, N; Vaccine 1995, V13, P1152 CAPLUS
(6) Clarke, B; Nature 1987, V330, P381 CAPLUS
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 8 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1999:479138 CAPLUS
DOCUMENT NUMBER: 132:11552
TITLE: T cell response pattern to glutamic acid decarboxylase
65 (GAD65) peptides of newly diagnosed type 1 diabetic
patients sharing susceptible HLA haplotypes
AUTHOR(S): Rharbaoui, F.; Mayer, A.; Granier, C.; Bouanani, M.;
Thivolet, C.; Pau, B.; Orgiazzi, J.; Madec, A.-M.
CORPORATE SOURCE: Faculte de Pharmacie, Montpellier, Fr.
SOURCE: Clin. Exp. Immunol. (1999), 117(1), 30-37
CODEN: CEXIAL; ISSN: 0009-9104
PUBLISHER: Blackwell Science Ltd.
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Autoantibodies and autoreactive T lymphocytes directed against several
pancreatic .beta. cell proteins such as GAD65 have been identified in the
circulation before and at the onset of clin. type 1 (insulin-dependent)
diabetes. Using GAD65 synthetic peptides, the authors studied the

proliferative response of peripheral blood mononuclear cells (PBMC) either from recently diagnosed type 1 diabetic patients, of whom the majority share the disease-assocd. HLA class II haplotype (DR4-DQB1*0201 or DR3-DQB1*0302), or from HLA-matched control subjects. The authors found that 67% (14/21) of the type 1 diabetic patients and 39% (9/23) of the control subjects exhibited a pos. proliferative response. Compared with control subjects, however, PBMC from diabetic patients proliferated more frequently in the presence of peptide pools from the C-terminal region of GAD65 (amino acids 379-585). Diabetic patients with the same HLA-DQ or HLA-DR alleles showed partially identical T cell reactivity, but no clear correlation could be made between MHC class II specificity and T cell epitopes because of multiple combinations of class II alleles. In addn., by flow cytometry, the authors studied the direct binding of GAD65 peptides to MHC class II mols. of Epstein-Barr virus (EBV)-transformed B (EBV-B) cells obtained from a diabetic patient. They found that 11 GAD peptides were able to bind to the highly susceptible haplotype DRB1*0301/0401-DQA1*0301/0501-DQB1*0302/0201 on the surface of EBV-B cells in partial correlation with the results obtained in the proliferation assays.

IT 251651-17-9

RL: PRP (Properties)

(T cell response pattern to glutamic acid decarboxylase 65 (GAD65) peptides of newly diagnosed type 1 diabetic patients sharing susceptible HLA haplotypes)

REFERENCE COUNT:

32

REFERENCE(S):

- (2) Atkinson, M; J Clin Invest 1994, V94, P2125 CAPLUS
- (4) Baekkeskov, S; Nature 1990, V347, P151 CAPLUS
- (7) Brooks-Worrell, B; J Immunol 1996, V157, P5668 CAPLUS
- (9) Calvo-Calle, J; J Immunol 1997, V159, P1362 CAPLUS
- (10) Durinovic-Bello, I; Diabetes 1996, V45, P795 CAPLUS

ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 9 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1999:393960 CAPLUS

DOCUMENT NUMBER: 131:39718

TITLE: Antifungal and antibacterial histatin-based peptides

INVENTOR(S): Oppenheim, Frank G.; Xu, Tao; Roberts, F. Donald; Spacciapoli, Peter; Friden, Phillip M.

PATENT ASSIGNEE(S): Periodontix, Inc., USA; Trustees of Boston University

SOURCE: U.S., 35 pp., Cont.-in-part of U.S. 5,631,228.

CODEN: USXXAM

DOCUMENT TYPE:

Patent

LANGUAGE:

English

FAMILY ACC. NUM. COUNT: 7

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5912230	A	19990615	US 1998-973559	19980311
US 5486503	A	19960123	US 1994-287717	19940809
US 5631228	A	19970520	US 1995-481888	19950607
WO 9640768	A2	19961219	WO 1996-US9374	19960607
WO 9640768	A3	19970306		

W: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG

RW: KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA

PRIORITY APPLN. INFO.:

US 1991-786571 B1 19911101
 US 1993-145030 B1 19931028
 US 1994-287717 A2 19940809
 US 1995-481888 A2 19950607
 WO 1996-US9374 W 19960607

AB Histatin-based peptides representing defined portions of the amino acid sequences of naturally occurring human histatins, and methods for treatment of fungal or bacterial infection, are described. The histatin-based peptides represent the active anti-fungal and anti-bacterial region of naturally occurring human histatins.

IT 186258-89-9 186258-89-9D, derivs.

RL: BAC (Biological activity or effector, except adverse); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(histatin-based peptides for antifungal and antibacterial agents)

REFERENCE COUNT:

41

REFERENCE(S):

- (1) Anon; JP 03261747 1994 CAPLUS
 - (2) Anon; JP 06234653 1994 CAPLUS
 - (3) Anon; JP 06287146 1994 CAPLUS
 - (4) Anon; WO 9421672 1994 CAPLUS
 - (5) Chang, C; Peptides 1990 1991, P843 CAPLUS
- ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 10 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1999:208580 CAPLUS

DOCUMENT NUMBER: 130:232464

TITLE: Antifungal and antibacterial D-amino acid histatin-based peptides

INVENTOR(S): Oppenheim, Frank G.; Xu, Tao; Roberts, F. Donald; Spacciapoli, Peter; Friden, Phillip M.

PATENT ASSIGNEE(S): Periodontix, Inc., USA; Trustees of Boston University

SOURCE: U.S., 34 pp., Cont.-in-part of U.S. 5,631,228.

CODEN: USXXAM

DOCUMENT TYPE:

Patent

LANGUAGE:

English

FAMILY ACC. NUM. COUNT: 7

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5885965	A	19990323	US 1998-973563	19980312
US 5486503	A	19960123	US 1994-287717	19940809
US 5646119	A	19970708	US 1995-485273	19950607
WO 9640770	A2	19961219	WO 1996-US9962	19960607
WO 9640770	A3	19970206		

W: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG

RW: KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA

PRIORITY APPLN. INFO.:

US 1991-786571 B1 19911101
 US 1993-145030 B1 19931028
 US 1994-287717 A2 19940809
 US 1995-485273 A2 19950607
 WO 1996-US9962 W 19960607

AB D-Amino acid histatins and histatin-based peptides and methods for treatment of fungal or bacterial infection are described. These D-amino acid histatins and histatin-based peptides are longer-acting antifungal or antibacterial agents than their L-enantiomeric analogs.

IT 186258-89-9D, D-amino acid-contg.

RL: BAC (Biological activity or effector, except adverse); PRP

(Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(antifungal and antibacterial D-amino acid histatin-based peptides)

REFERENCE COUNT: 32

REFERENCE(S): (1) Anon; JP 03261747 1994 CAPLUS
(2) Anon; JP 06234653 1994 CAPLUS
(3) Anon; JP 06287146 1994 CAPLUS
(4) Anon; WO 9421672 1994 CAPLUS
(5) Berkowitz; US 5459237 1995 CAPLUS
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 11 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1999:169524 CAPLUS

DOCUMENT NUMBER: 131:1843

TITLE: A 7-amino-acid insert in the heavy chain nucleotide binding loop alters the kinetics of smooth muscle myosin in the laser trap

AUTHOR(S): Lauzon, Anne-Marie; Tyska, Matthew J.; Rovner, Arthur S.; Freydon, Yelena; Warshaw, David M.; Trybus, Kathleen M.

CORPORATE SOURCE: Department of Molecular Physiology and Biophysics, University of Vermont, Burlington, VT, 05405, USA

SOURCE: J. Muscle Res. Cell Motil. (1998), 19(8), 825-837
CODEN: JMRMD3; ISSN: 0142-4319

PUBLISHER: Kluwer Academic Publishers

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Two smooth muscle myosin heavy chain isoforms differ by a 7-amino-acid insert in a flexible surface loop located near the nucleotide binding site. The non-inserted isoform is predominantly found in tonic muscle, while the inserted isoform is mainly found in phasic muscle. The inserted isoform has twice the actin-activated ATPase activity and actin filament velocity in the in vitro motility assay as the non-inserted isoform. We used the laser trap to characterize the mol. mechanics and kinetics of the inserted isoform ((+)insert) and of a mutant lacking the insert ((-)insert), analogous to the isoform found in tonic muscle. The constructs were expressed as heavy meromyosin using the baculovirus/insect cell system. Unitary displacement (d) was similar for both constructs (.apprx. 10 nm) but the attachment time τ_{on} for the (-)insert was twice as long as for the (+)insert regardless of the [MgATP]. Both the relative av. isometric force ($F_{avg}(-insert)/F_{avg}(+insert) = 1.1 \pm 0.2$ (mean \pm SE) using the in vitro motility mixt. assay, and the unitary force (F .apprx. 1 pN) using the laser trap, showed no difference between the two constructs. However, as under unloaded conditions, τ_{on} under loaded conditions was longer for the (-)insert compared with the (+)insert construct at limiting [MgATP]. These data suggest that the insert in this surface loop does not affect the mechanics but rather the kinetics of the cross-bridge cycle. Through comparisons of τ_{on} from d measurements at various [MgATP], we conclude that the insert affects two specific steps in the cross-bridge cycle, i.e., MgADP release and MgATP binding.

IT 225518-96-7

RL: BSU (Biological study, unclassified); BIOL (Biological study)
(7-amino-acid insert in heavy chain nucleotide binding loop alters kinetics of smooth muscle myosin in laser trap)

REFERENCE COUNT: 44

REFERENCE(S): (1) Babij, P; Nucleic Acid Res 1993, V21, P1467 CAPLUS
(2) Bobkov, A; Proc Natl Acad Sci USA 1996, V93, P2285 CAPLUS
(3) Cooke, R; Physiol Rev 1997, V77, P671 CAPLUS
(4) Cremonesi, C; Biochem 1998, V37, P1969 CAPLUS
(5) Cuda, G; Biophys J 1997, V72, P1767 CAPLUS
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L4 ANSWER 12 OF 25 CAPLUS COPYRIGHT 2001 ACS
 ACCESSION NUMBER: 1998:65926 CAPLUS
 DOCUMENT NUMBER: 128:127089
 TITLE: Methods for determining the presence of brain protein S-100
 INVENTOR(S): Brundell, Jan; Nyberg, Lena
 PATENT ASSIGNEE(S): AB Sangtec Medical, Swed.; Brundell, Jan; Nyberg, Lena
 SOURCE: PCT Int. Appl., 29 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9801471	A1	19980115	WO 1997-SE1164	19970627
W: AU, BR, CA, FI, HU, IL, JP, NO, NZ, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
CA 2259413	AA	19980115	CA 1997-2259413	19970627
AU 9735633	A1	19980202	AU 1997-35633	19970627
AU 715797	B2	20000210		
EP 931094	A1	19990728	EP 1997-932092	19970627
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
BR 9710175	A	20000111	BR 1997-10175	19970627
JP 2000515854	T2	20001128	JP 1998-505124	19970627
NO 9806218	A	19981230	NO 1998-6218	19981230
PRIORITY APPLN. INFO.:			SE 1996-2677	A 19960705
			WO 1997-SE1164	W 19970627

AB Assay methods for detg. the presence of the brain protein S-100 in a clin. sample which use antibodies directed to epitopes in the region from Ser1 to Asn38 and from Thr82 to Glu93 of the amino acid sequence of the .beta. subunit of human S100B is provided. The methods developed include luminescent immunoassay, ELISA, and IRMA, and is used for detecting cerebral dysfunction, melanoma cancer, and for evaluating the influence of extracorporeal circulation equipment on injured brain.

IT 201799-94-2 201799-95-3
 RL: BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
 (monoclonal antibodies to human brain protein S-100 epitopes for diagnosis of cerebral dysfunction and melanoma)

L4 ANSWER 13 OF 25 CAPLUS COPYRIGHT 2001 ACS
 ACCESSION NUMBER: 1998:62220 CAPLUS
 DOCUMENT NUMBER: 128:136508
 TITLE: Peptide inhibitors of selectin binding, preparation, and therapeutic and diagnostic use
 INVENTOR(S): Heavner, George A.; Kruszynski, Marian
 PATENT ASSIGNEE(S): Centocor, Inc., USA
 SOURCE: U.S., 29 pp. Cont.-in-part of U.S. Ser. No. 997,771, abandoned.
 CODEN: USXXAM
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 2
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
-----	-----	-----	-----	-----

US 5710123 A 19980120 US 1995-454207 19950609
 WO 9414836 A1 19940707 WO 1993-US12110 19931213
 W: CA, JP, US
 RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE
 PRIORITY APPLN. INFO.: US 1992-997771 19921218
 WO 1993-US12110 19931213
 OTHER SOURCE(S): MARPAT 128:136508
 AB The present invention provides novel peptides having as their core region portions of the 109-118 amino acid sequence of P-selectin, E-selectin or L-selectin. The invention also provides pharmaceutical compns. comprising the peptides of the invention, and diagnostic and therapeutic methods utilizing the peptides and pharmaceutical compns. of the invention.
 IT **161174-87-4P 161174-90-9P**
 RL: BAC (Biological activity or effector, except adverse); PRP (Properties); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
 (peptide inhibitors of selectin binding, prepn., and therapeutic and diagnostic use)
 L4 ANSWER 14 OF 25 CAPLUS COPYRIGHT 2001 ACS
 ACCESSION NUMBER: 1997:244274 CAPLUS
 DOCUMENT NUMBER: 126:224281
 TITLE: cDNAs for the plant panallergen co-factor-independent phosphoglycerate mutase and identification of diagnostic and therapeutically useful epitopes
 INVENTOR(S): Ferreira, Fatima; Richter, Klaus; Engel, Edwin; Ebner, Christof; Jilek, Alexander; Rheinberger, Hans-Joerg; Kraft, Dietrich; Breitenbach, Michael
 PATENT ASSIGNEE(S): Biomay Produktions- Und Handelsgesellschaft MbH, Austria; Ferreira, Fatima; Richter, Klaus; Engel, Edwin; Ebner, Christof; Jilek, Alexander; Rheinberger, Hans-Joerg; Kraft, Dietrich; Breitenbach, Michael
 SOURCE: PCT Int. Appl., 159 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: German
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9705258	A2	19970213	WO 1996-AT141	19960802
WO 9705258	A3	19970327		
W: AU, CA, JP, NO, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
AT 9501320	A	19961015	AT 1995-1320	19950802
AT 402505	B	19970625		
AU 9666059	A1	19970226	AU 1996-66059	19960802
PRIORITY APPLN. INFO.:			AT 1995-1320	19950802
			WO 1996-AT141	19960802
AB CDNAs for the pollen panallergen co-factor-independent phosphoglycerate mutase (E.C. 5.4.2.1.) of birch, mugwort and timothy grass pollen are cloned and characterized. This sequence of the allergen is highly conserved in all plants, but not in animals. The amino acid sequence and the most important B and T cell epitopes of the mol. are derived and demonstrated. The allergen was manufd. in E. coli and bound the IgE serum of patients who are allergic to tree, grass and weed pollens and various foodstuffs. A monoclonal antibody (BIP 3) specifically binds to this protein from all plants tested. The significance of the co-factor-independent phosphoglycerate mutase (E.C. 5.4.2.1.) derives from the fact that it results in the cross-sensitization of patients. The				

protein and peptide fragments can be used in diagnostic and therapeutic methods based, for example, on antigen -antibody interaction, mediator release or T-cell reactivity.

IT 187816-25-7

RL: ADV (Adverse effect, including toxicity); BPR (Biological process); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)

(B-cell epitope of mugwort pollen phosphoglycerate mutase isoenzyme Art17; cDNAs for plant panallergen co-factor-independent phosphoglycerate mutase and identification of diagnostic and therapeutically useful epitopes)

IT 187816-07-5

RL: ADV (Adverse effect, including toxicity); BPR (Biological process); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)

(B-cell epitope of mugwort pollen phosphoglycerate mutase isoenzyme Art6; cDNAs for plant panallergen co-factor-independent phosphoglycerate mutase and identification of diagnostic and therapeutically useful epitopes)

L4 ANSWER 15 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1997:116549 CAPLUS

DOCUMENT NUMBER: 126:113162

TITLE: Anti-fungal and anti-bacterial histatin-based peptides

INVENTOR(S): Oppenheim, Frank G.; Xu, Tao; Roberts, F. Donald;

Spacciapoli, Peter; Friden, Phillip M.

PATENT ASSIGNEE(S): Periodontix, Inc., USA; Trustees of Boston University;

Oppenheim, Frank G.; Xu, Tao; Roberts, F. Donald;

Spacciapoli, Peter; Friden, Phillip M.

SOURCE: PCT Int. Appl., 70 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 7

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9640768	A2	19961219	WO 1996-US9374	19960607
WO 9640768	A3	19970306		
W:	AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG			
RW:	KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA			
US 5631228	A	19970520	US 1995-481888	19950607
AU 9661585	A1	19961230	AU 1996-61585	19960607
AU 709204	B2	19990826		
EP 832119	A2	19980401	EP 1996-919182	19960607
EP 832119	B1	20000920		
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			
JP 11508238	T2	19990721	JP 1996-501707	19960607
AT 196475	E	20001015	AT 1996-919182	19960607
US 5912230	A	19990615	US 1998-973559	19980311
PRIORITY APPLN. INFO.:			US 1995-481888	A 19950607
			US 1991-786571	B1 19911101
			US 1993-145030	B1 19931028
			US 1994-287717	A2 19940809
			WO 1996-US9374	W 19960607

OTHER SOURCE(S): MARPAT 126:113162

AB Histatin-based peptides representing defined portions of the amino acid sequences of naturally occurring human histatins and methods for treatment of fungal or bacterial infection are described. These histatin-based peptides represent the active anti-fungal and anti-bacterial region of naturally occurring human histatins.

IT 186258-89-9

RL: BAC (Biological activity or effector, except adverse); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses) (anti-fungal and anti-bacterial histatin-based peptides)

L4 ANSWER 16 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1996:260618 CAPLUS

DOCUMENT NUMBER: 124:314670

TITLE: A protective anti-peptide antibody against the immunodominant site of the A24 Cruzeiro strain of foot-and-mouth disease virus and its reactivity with other subtype viruses containing the same minimum binding sequence

AUTHOR(S): Barnett, P. V.; Pullen, L.; Staple, R. F.; Lee, L. J.; Butcher, R.; Parkinson, D.; Doel, T. R.

CORPORATE SOURCE: Inst. Animal Health, Pirbright Lab., Pirbright, Woking, Surrey, GU24 0NF, UK

SOURCE: J. Gen. Virol. (1996), 77(5), 1011-18
CODEN: JGVIAY; ISSN: 0022-1317

DOCUMENT TYPE: Journal

LANGUAGE: English

AB A synthetic peptide vaccine of the general sequence Cys-Cys-(200-213)-Pro-Pro-Ser-(141-158)-Pro-Cys-Gly(peptide A40), where the numbered residues refer to the VP1 sequence of foot-and-mouth disease virus (FMDV) strain A24 Cruzeiro, has previously been shown to elicit neutralizing and protective antibodies in guinea-pigs and cattle. To examine this immunogenic tract in more detail monoclonal antibodies (MAbs) were raised to this peptide. One such MAb, Cl.1, which recognized the homologous peptide, bound to native virus, neutralized infectivity in vitro and passively protected mice from challenge. Using overlapping dodecameric peptides the min. binding 'footprint' of this MAb incorporated residues 149-154 which were resp. Gly-Ser-Leu-Ala-Ala-Arg. Since this 'footprint' occurs in several other A subtype strains of FMDV, the extent to which MAb Cl.1 could cross-react was also examd. Using a liq.-phase competition ELISA, only viruses with a sequence that encompassed the same min. binding 'footprint', namely A27 Cundinamarca Colombia/76, A Argentina/79, and A Venceslau Brazil/76 reacted with similar affinity against MAb Cl.1. However, further serol. examn. of Cl.1 with these viruses by indirect ELISA, in vitro neutralization and passive protection showed clear functional disparity. In contrast to the liq.-phase ELISA, the ability of Cl.1 to react with electrostatically bound virus varied significantly depending on the subtype examd. Moreover, the capacity of this MAb to neutralize these subtypes showed wide divergence which was mirrored by the protection data.

IT 176260-59-6 176260-60-9

RL: BPR (Biological process); PRP (Properties); BIOL (Biological study); PROC (Process)
(epitope mapping and neutralizing activity of anti-VP1 monoclonal antibody against peptide of foot-and-mouth disease virus A24 Cruzeiro)

L4 ANSWER 17 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1995:380112 CAPLUS

DOCUMENT NUMBER: 122:161381

TITLE: Preparation of peptide inhibitors of selectin binding.

INVENTOR(S): Heavner, George A.; Kruszynski, Marian

PATENT ASSIGNEE(S): Centocor, Inc., USA
 SOURCE: PCT Int. Appl., 52 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 2
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9414836	A1	19940707	WO 1993-US12110	19931213
W: CA, JP, US				
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
US 5710123	A	19980120	US 1995-454207	19950609
PRIORITY APPLN. INFO.:			US 1992-997771	19921218
			WO 1993-US12110	19931213

OTHER SOURCE(S): MARPAT 122:161381

AB R1X1A1B1C1D1E1F1G1H1I1J1X2R2, R1X1-cyclo-(A2B1C1D1E1F1G1H1I2)-J1X2R2 (X1 = N-terminal sequence of 0-10 amino acids; R1 = H, alkyl, aryl, CHO, alkanoyl, aroyl, alkoxycarbonyl, aryloxycarbonyl; X2 = C-terminal sequence of 0-10 amino acids; R2 = OH, OR3, NR5R6; R3 = alkyl, aryl; R5, R6 = H, alkyl, aryl, cycloalkyl; A1 = null, D- or L-Cys; A2 = D- or L-Cys; B1 = D- or L-His, -Ser, -Leu, -Phe, -Asn, -Pro, -Glu; C1 = D- or L-Lys, -His, -Arg, -Ser; D1 = D- or L-Lys, -Leu, -Ala, -Phe, -His, -Arg, -Ser; E1 = D- or L-Lys, -Phe, -Gln, -Arg; F1 = D- or L-His, -Leu, -Ala, -Ile, -Thr, -Arg; G1 = D- or L-Ala, -Phe, -His, -Gln; H1 = D- or L-Leu, -Phe, -Ile, -Pro, -Ala; I1 = D- or L-Cys, -Phe, -Ile, -His, -Leu, -Val, -Thr, -Ser; I2 = D- or L-Cys; J1 = D- or L-Tyr, -Phe, -Ile, -Val), were prepd. The peptides have as their core region portions of the 109-118 amino acid sequence of P-selectin, E-selectin or L-selectin. Diagnostic and therapeutic methods are given utilizing the peptides and pharmaceutical compns. thereof. Thus, Cys-Leu-Lys-Lys-Lys-His-Ala-Leu-Cys-Tyr-NH2 was prepd. using BOC-protected amino acids on methylbenzhydrylamine resin. Title compds. inhibited binding of human neutrophils to P-selectin with IC50 = 0.003-1.013 mM.

IT 161174-87-4P 161174-90-9P

RL: BAC (Biological activity or effector, except adverse); BUU (Biological use, unclassified); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
 (prepn. of peptide inhibitors of selectin binding)

L4 ANSWER 18 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1994:555572 CAPLUS

DOCUMENT NUMBER: 121:155572

TITLE: Experimental autoimmune panencephalitis and uveoretinitis transferred to the Lewis rat by T lymphocytes specific for the S100.beta. molecule, a calcium binding protein of astroglia

AUTHOR(S): Kojima, Kimikazu; Berger, Thomas; Lassmann, Hans; Hinze-Selch, Dunja; Zhang, Yiping; Gehrmann, Jochen; Reske, Konrad; Wekerle, Hartmut; Linington, Christopher

CORPORATE SOURCE: Dep. Neuroimmunol., Max-Planck Inst. Psychiatr., Martinsried, 82152, Germany

SOURCE: J. Exp. Med. (1994), 180(3), 817-29

CODEN: JEMEAV; ISSN: 0022-1007

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The pathogenic potential of autoimmune T cell responses to nonmyelin autoantigens was investigated in the Lewis rat using the astrocyte-derived calcium binding protein S100.beta., as a model nonmyelin autoantigen. The

Lewis rat mounts a vigorous RT1B (major histocompatibility complex class II) restricted autoimmune response to an immunodominant S100.beta. epitope (amino acid residues 76-91). The adoptive transfer of S100.beta.-specific T cell lines induced a severe inflammatory response in the nervous system, but only minimal neurol. dysfunction in naive syngeneic recipients. The inability of S100.beta.-specific T cell transfer to induce severe disease was assocd. with a decreased recruitment of ED1+ macrophages into the central nervous system (CNS) in comparison with that seen in severe exptl. autoimmune encephalomyelitis (EAE) induced by the adoptive transfer of myelin basic protein (MBP)-specific T line cells. Moreover, unlike encephalitogenic MBP-specific T lines, S100.beta.-specific T cell lines exhibited no cytotoxic activity in vitro. Histopathol. anal. also revealed striking differences in the distribution of inflammatory lesions in MBP- and S100.beta.-specific T cell-mediated disease. In contrast to the MBP paradigm, S100.beta.-specific T cell transfer induces intense inflammation not only in the spinal cord, but throughout the entire CNS and also in the uvea and retina of the eye. In view of the distribution of lesions throughout the gray and white matter of the CNS the authors propose to term this new model exptl. autoimmune panencephalomyelitis (EAP) to differentiate it from EAE. These expts. demonstrate for the first time that nonmyelin CNS autoantigens can initiate a pathogenic autoimmune T cell response, although the nature of the target autoantigen profoundly influences the clin. and histopathol. characteristics of the resulting autoimmune disease. This is not simply a consequence of the distribution of the autoantigen, as both MBP and S100.beta. are coexpressed in many areas of the CNS, but reflects differences in the capacity of different regions of the CNS to process and present autoantigens. This new model of T cell-mediated autoimmune CNS disease exhibits a no. of similarities to multiple sclerosis (MS), such as its mild clin. course and the involvement of areas of the brain and eye, which are absent in myelin-mediated models of EAE. Nonmyelin autoantigens may therefore play an unexpectedly important role in the immunopathogenesis of inflammatory diseases of the CNS.

IT 157566-35-3

RL: BIOL (Biological study)

(in epitope mapping for nonmyelin autoantigen S100.beta., autoimmune panencephalitis and uveoretinitis in relation to)

L4 ANSWER 19 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1992:658207 CAPLUS

DOCUMENT NUMBER: 117:258207

TITLE: Antiasthmatic compositions containing peptides

INVENTOR(S): Nishimoto, Ikuo; Okamoto, Taku; Okuni, Yoshihiro

PATENT ASSIGNEE(S): Japan

SOURCE: Jpn. Kokai Tokkyo Koho, 5 pp.

CODEN: JKXXAF

DOCUMENT TYPE: Patent

LANGUAGE: Japanese

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
JP 04164036	A2	19920609	JP 1990-285748	19901025

AB Antiasthmatic comps. contain peptides having basic amino acids (arginine, lysine, and histidine) at the N-terminal 1st and 2nd positions and the C-terminal last position and the 3rd or 4th positions therefrom. For example, 9 peptides are presented, but no pharmacol. data given.

IT 144776-40-9 144776-41-0 144776-43-2

RL: BIOL (Biological study)

(as antiasthmatic)

L4 ANSWER 20 OF 25 CAPLUS COPYRIGHT 2001 ACS
 ACCESSION NUMBER: 1989:534755 CAPLUS
 DOCUMENT NUMBER: 111:134755
 TITLE: Preparation of decapeptides as LHRH antagonists having high antioviulatory activity and negligible histamine releasing activity
 INVENTOR(S): Folkers, Karl; Bowers, Cyril Y.; Ljungquist, Anders; Tang, Pui Fun Louisa; Kobota, Minoru; Feng, Dong Mei
 PATENT ASSIGNEE(S): University of Texas System, USA
 SOURCE: PCT Int. Appl., 70 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 2
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 8901944	A1	19890309	WO 1988-US2922	19880824
W: AT, AU, BB, BG, BR, CH, DE, DK, FI, GB, HU, JP, KP, KR, LK, LU, MC, MG, MW, NL, NO, RO, SD, SE, SU, US				
RW: AT, BE, BJ, CF, CG, CH, CM, DE, FR, GA, GB, IT, LU, ML, MR, NL, SE, SN, TD, TG				
US 4935491	A	19900619	US 1987-88431	19870824
AU 8825294	A1	19890331	AU 1988-25294	19880824
AU 619221	B2	19920123		
EP 377665	A1	19900718	EP 1988-908786	19880824
EP 377665	B1	19950712		
R: AT, BE, CH, DE, FR, GB, IT, LI, LU, NL, SE				
JP 03501969	T2	19910509	JP 1988-507982	19880824
HU 59940	A2	19920728	HU 1988-5868	19880824
HU 213098	B	19970228		
CA 1339659	A1	19980203	CA 1988-587364	19881230
DK 9000486	A	19900419	DK 1990-486	19900223
NO 9000888	A	19900423	NO 1990-888	19900223
NO 9402179	A	19900423	NO 1994-2179	19940610
PRIORITY APPLN. INFO.:			US 1987-88431	19870824
			WO 1988-US2922	19880824
			NO 1990-888	19900223
AB	Decapeptide analogs of LHRH, e.g. [N-Ac-D-2-Nal1, D-pClPhe2, D-3-Pal3, NicLys5, D-NicLys6, Ilys8, D-Ala10]-LHRH [2-Nal = 3-(2-naphthyl)alanine, pClPhe = 3-(4-chloro)phenylalanine, 3-Pal = 3-(3-pyridyl)alanine, NicLys = N.epsilon.-nisotinoyllysine, Ilys = N.epsilon.-isopropyllysine] (I) (Antide) having high ovulation inhibition activity and very low histamine release activity, were prepd. I and other decapeptides were synthesized by the solid phase method using a Beckman Model 990 peptide synthesizer, new lysine, ornithine, alanine, glutamic acid and arginine derivs., and benzhydrylamine hydrochloride resin as a solid support. I showed antioviulatory activity (AOA) of 100% at 1 .mu.g and 36% at 0.5 .mu.g in rats and an ED50 of .gtoreq.300 .mu.g/mL for histamine release in a rat mast cell assay.			
IT	112424-93-8P RL: SPN (Synthetic preparation); PREP (Preparation) (prepn. of, as LHRH antagonist)			

L4 ANSWER 21 OF 25 CAPLUS COPYRIGHT 2001 ACS
 ACCESSION NUMBER: 1989:455372 CAPLUS
 DOCUMENT NUMBER: 111:55372
 TITLE: Neutralizing epitopes of type O foot-and-mouth disease virus. II.. Mapping three conformational sites with

synthetic peptide reagents
AUTHOR(S): Parry, N. R.; Barnett, P. V.; Ouldrige, E. J.;
Rowlands, D. J.; Brown, F.
CORPORATE SOURCE: Dep. Virol., Wellcome Biotechnol. Ltd.,
Beckenham/Kent, BR3 3BS, UK
SOURCE: J. Gen. Virol. (1989), 70(6), 1493-503
CODEN: JGVIAI; ISSN: 0022-1317
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Four neutralizing monoclonal antibodies (MAbs), recognizing 3 functionally independent, conformational sites on type O foot-and-mouth disease virus (FMDV) failed to react with immobilized structural proteins or synthetic peptides but bound to the isolated capsid protein VP1 and peptides in soln. Inhibition ELISA techniques were, therefore, applied using peptide antigens and anti-peptide sera to block MAb binding to virus particles, permitting the identification of those portions of the VP1 protein contributing to the epitopes. The binding site of one MAb, which neutralized a range of type O FMDV isolates, was shown to have components within regions 146-150 and 200-213 of VP1 with a crit. involvement of the amino acids at positions 146 and 206 or 207. The determinants recognized by 2 other MAbs which were directed at similar, but not identical, epitopes from a second site included components from the 200 to 213 and 143 to 146 regions with amino acids 143 and 144, resp., being crit. for the inhibition of the virus binding of the 2 antibodies. Thus, the 2 previously identified immunogenic tracts of VP1 are brought into proximity in the quaternary structure of the virion to form an antigenic domain contg. several conformational epitopes, some of which are functionally independent. A fourth, strain-specific MAb was effectively blocked from reacting with virus by peptides corresponding to residues 161-180 and 200-213.
IT 121795-39-9
RL: PROC (Process)
(monoclonal antibody recognition of, of foot-and-mouth disease virus VP1 protein)

L4 ANSWER 22 OF 25 CAPLUS COPYRIGHT 2001 ACS
ACCESSION NUMBER: 1988:433575 CAPLUS
DOCUMENT NUMBER: 109:33575
TITLE: Comparison of benzo[a]pyrene-diol-epoxide binding to histone H2A with different carboxy-terminal regions
AUTHOR(S): Kurokawa, Masahiko; MacLeod, Michael C.
CORPORATE SOURCE: Syst. Cancer Cent., Univ. Texas, Smithville, TX, USA
SOURCE: Carcinogenesis (London) (1988), 9(3), 419-25
CODEN: CRNGDP; ISSN: 0143-3334
DOCUMENT TYPE: Journal
LANGUAGE: English
AB (.-.-)-7r,8t-Dihydroxy-9t,10t-oxy-7,8,9,10-tetrahydrobenzo[a]pyrene (DPBE-I) binding was compared in 3 vertebrate histones H2A with different C-terminal regions. HPLC analyses of core histones prep'd. from nuclei exposed to [3H]BPDE-I showed that rat liver and chicken erythrocyte histones H2A were heavily labeled by [3H]BPDE-I, but Xenopus laevis liver histone H2A was not. This result was confirmed by HPLC analyses of V8-protease digests of BPDE-I bound to histone H2A purified from the 3 different nuclei. There are significant amino acid sequence differences only in the C-terminal regions of the different histones H2A, where rat liver and chicken erythrocyte histones H2A contain 2 and 1 histidine residues, resp., while the amino acid sequence of Xenopus histone H2A contains no histidine. Pre-treatment of the in situ BPDE-I-modified H2A.2 from rat liver with carboxypeptidase B, which should remove the C-terminal lysine from the protein, resulted in increased retention times on reverse-phase HPLC for the adduct-contg. peptides upon subsequent

V8-protease digestion. This suggested that the site(s) of BPDE-I modification are located primarily in the C-terminal octapeptide of rat H2A.2. To confirm this, C-terminal V8-peptides of the different histones H2A were isolated and reacted with BPDE-I at physiol. pH in vitro. The HPLC analyses of the reaction mixts. indicated that the C-terminal peptide of rat liver and chicken erythrocyte histones H2A was a target site for BPDE-I binding in nuclei. Thus, the nucleophilic target amino acid for BPDE-I binding in histone H2A may be a histidine located close to the C terminus.

IT 115154-41-1 115154-42-2

RL: BIOL (Biological study)

(of C-terminal of histone H2A, benzopyrene diol epoxide binding in relation to, species differences in)

L4 ANSWER 23 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1988:49436 CAPLUS

DOCUMENT NUMBER: 108:49436

TITLE: Design, synthesis and bioassays of antagonists of LH-RH which have high antioviulatory activity and release negligible histamine

AUTHOR(S): Ljungqvist, Anders; Feng, Dong Mei; Tang, Pui Fun
Louisa; Kubota, Minoru; Okamoto, Tadashi; Zhang, Yawen; Bowers, Cyril Y.; Hook, William A.; Folkers, Karl

CORPORATE SOURCE: Inst. Biomed. Res., Univ. Texas, Austin, TX, 78712, USA

SOURCE: Biochem. Biophys. Res. Commun. (1987), 148(2), 849-56
CODEN: BBRCA9; ISSN: 0006-291X

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Potent antagonists of LH-RH were achieved which release negligible histamine. [N-Ac-D-2-Nal1,D-pClPhe2,D-3-Pal3,NicLys5,D-NicLys6,ILys8,D-Ala10]-LHRH (2-Nal = 3-(2-naphthylalanine; 3-Pal = 3-(3-pyridylalanine); NicLys = N.epsilon.-nicotinoyllysine; ILys = N.epsilon.-isopropyllysine) showed 100% antioviulatory activity (AOA)/1 .mu.g and 36% AOA/0.5 .mu.g; the ED50 for histamine release was >300 .mu.g/mL. [N-Ac-D-2-Nal1,D-pClPhe2,D-3-Pal3,PicLys5,D-PicLys6,ILys8,D-Ala10]-LHRH (PicLys = N.epsilon.-picoloyllysine) showed 100% AOA/0.5 .mu.g, 40% AOA/0.25 .mu.g, and an ED50 for histamine release of 93 .mu.g/mL; it was the most potent of 52 new peptides. These antagonists feature designs with weakly basic acylated D-Lys6 (D-NicLys6 and D-PicLys6), alkylated Lys8 or Orn8 (ILys8 and IOrn8 (N.delta.-isopropylornithine)), NicLys5, and PicLys5. Concepts included balanced overall basicity, superiority of ILys8 and IOrn8 which are sequence dependent, and sensitivity of positions 5 and 6 for potency.

IT 112424-93-8

RL: BIOL (Biological study)

(ovulation inhibition by, with negligible histamine release, mol. structure in relation to)

L4 ANSWER 24 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1985:465120 CAPLUS

DOCUMENT NUMBER: 103:65120

TITLE: Products from catabolism of proteins as analogs of peptide bioregulators (structural factors of the physiological activity)

AUTHOR(S): Leonova, V. I.; Tseitlin, V. M.; Galaktionov, S. G.; Nikolaichik, V. V.

CORPORATE SOURCE: Opornyi Punkt, VNII Genet. Sel. Prom. Mikroorg., Minsk, USSR

SOURCE: Konform. Funkts. Biol. Mol. (1984), 73-81. Editor(s):

Chipens, G. I. Zinatne: Riga, USSR.
CODEN: 53VVAG

DOCUMENT TYPE: Conference
LANGUAGE: Russian

AB The presence of sequences of peptide bioregulators' fragments in blood proteins and their potential release following degrdn. of the proteins by enzymes present in circulating blood were discussed. E.g., the tuftsin analog His,Lys,Pro [97461-89-7] was found in fragment 535-538 of serum albumin and could be released by plasmin. The enkephalin analog Tyr-Gly-Gly-Phe-Tyr [97461-90-0] could be released from the .beta.-chain of fibrinogen by plasmin, esterase, and exopeptidase. The angiotensin analogs Glu-Glys-Ile-Tyr-Ile-His-Pro-Arg [97474-91-4] and Glu-Lys-Ile-Tyr-Ile-His-Pro-Arg-Tyr [97461-88-6] were released from fragments 405-412 and 405-413 of prothrombin [9001-26-7]. The stabilities of the conformational analogs of these bioregulator peptides were given and related to potential bioacty.

IT 97461-89-7

RL: BIOL (Biological study)
(release from serum albumin and stability of)

L4 ANSWER 25 OF 25 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 1985:94015 CAPLUS

DOCUMENT NUMBER: 102:94015

TITLE: Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid
AUTHOR(S): Geysen, H. Mario; Meloen, Rob H.; Barteling, Simon J.
CORPORATE SOURCE: Commonw. Serum Lab., Parkville, 3052, Australia
SOURCE: Proc. Natl. Acad. Sci. U. S. A. (1984), 81(13), 3998-4002

CODEN: PNASA6; ISSN: 0027-8424

DOCUMENT TYPE: Journal
LANGUAGE: English

AB A procedure is described for rapid concurrent synthesis on solid supports of hundreds of peptides, of sufficient purity to react in an ELISA. Interaction of synthesized peptides with antibodies is then easily detected without removing them from the support. In this manner, an immunogenic epitope of the immunol. important coat protein of foot-and-mouth disease virus (type O1) was located with a resoln. of 7 amino acids, corresponding to amino acids 146-152 of that protein. Then, a complete replacement set of peptides in which all 20 amino acids were substituted in turn at every position within the epitope was synthesized, and the particular amino acids conferring specificity for the reaction with antibody were detd. The leucine residues at positions 148 and 151 were essential for reaction with antisera raised against intact virus. A lesser contribution was derived from the glutamine and alanine residues at positions 149 and 152, resp. Aside from the practical significance for locating and examg. epitopes at high resoln., these findings may lead to better understanding of the basis of antigen-antibody interaction and antibody specificity.

IT 92227-36-6DP, polymer-bound 92269-39-1DP, polymer-bound
92269-40-4DP, polymer-bound

RL: BAC (Biological activity or effector, except adverse); BIOL
(Biological study); PREP (Preparation)
(prepn. and antigenicity of)

FILE 'HOME' ENTERED AT 08:55:24 ON 06 JUL 2001


```

OS Quercus robur (English oak).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fagales; Fagaceae; Quercus.
OX NCBI_TaxID=38942;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-IN VITRO SHOOT CULTURES;
RA Gil B., Pastoriza E.M., Ballester A., Sanchez C.;
RT "Identification of a phase-change related mRNA in oak shoot cultures
derived from basal sprouts and crown branches.";
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ271778; CAB72442.1; -.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 79 AA; 8414 MW; 8E45CABF40F00B6F CRC64;

Query Match
Best Local Similarity 53.1%; Score 52; DB 10; Length 79;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 HGHGKHKNNKGK 13
DB 54 HGHGHHGKPG 65

RESULT 13
Q9FH30 PRELIMINARY; PRT; 333 AA.
AC Q9FH30;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GB|AF34232.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
features of the regions of 4,504,864 bp covered by sixty P1 and TAC
clones.";
RL DNA Res. 7:131-135(2000).
DR EMBL; AB022215; BAB1274.1; -.
SQ SEQUENCE 333 AA; 37450 MW; 86C21350D8B3089 CRC64;

Query Match
Best Local Similarity 53.1%; Score 52; DB 10; Length 333;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 HGHGKHKNNKGK 15
DB 128 HGHGSPHKHDKE 141

RESULT 14
Q9WAC1 PRELIMINARY; PRT; 554 AA.
AC Q9WAC1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```

```

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG15784 PROTEIN.
GN CG15784.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Anthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutcliffe G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spletter E., Spradling A.C., Stapleton M., Strung R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003434; AAP46035.1; -.
DR FlyBase; FBgn0029766; CG15784.
DR InterPro; IPR002395; -.
DR PRINTS; PR00334; KININOGEN.
SQ SEQUENCE 554 AA; 62329 MW; 9CE2F80A7A1D902D CRC64;

Query Match
Best Local Similarity 52.6%; Score 51.5; DB 5; Length 554;
Matches 9; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 KHGSG--HGKHKNNKGK 14
DB 201 KHGSGPHNHGRRGHR 217

RESULT 15
Q9VHR3 PRELIMINARY; PRT; 819 AA.
AC Q9VHR3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

```

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
DR EMBL: AE003454; AAF46744.1; -
DR FlyBase: FBgn0034645; CG10320.
SQ SEQUENCE 110 AA; 12091 MW; 47552EC0EC5E9D76 CRC64;

Query Match 54.1%; Score 53; DB 5; Length 110;
Best Local Similarity 66.7%; Pred. No. 0.45;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 HGHGCHKHKN 13
DB 97 HGHGHEHGDGK 108

RESULT 10

O9U412 PRELIMINARY; PRT; 3469 AA.
AC O9U412;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE SANT DOMAIN PROTEIN SMRTR.
GN SMR OR SMRTR OR CG4013.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94417957; PubMed=1048333;
RA Tsai C.-C., Kao H.-Y., Yao T.-P., McKee M., Evans R.M.,
RT "SMRTR, a *Drosophila* nuclear receptor coregulator, reveals that ECR-
mediated repression is critical for development."
RL Mol. Cell 4:1175-1186(1999).
DR EMBL: AF175223; AAD52614.1; -
DR FlyBase: FBgn0024308; Smr.
DR InterPro: IPR001004; -
DR InterPro: IPR001005; -
DR InterPro: IPR002086; -
DR Pfam: PF00249; myb_DNA-binding; 1.
DR PRINTS: PR00308; ANTIREFEZET.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN.1.
DR SMART: SM00395; SANT; 1.
SQ SEQUENCE 3469 AA; 36415 MW; 6284E1AC5C247CD9 CRC64;

Query Match 54.1%; Score 53; DB 5; Length 3469;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HGHGCHKHKN 11
DB 1790 HGHGCHKHKN 1799

RESULT 11

O9YVJ9 PRELIMINARY; PRT; 3502 AA.
AC O9YVJ9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE SMR PROTEIN.
GN SMR OR CG4013.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abail J.F., Abpayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktiroglu I., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandal D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mel M.-H., Ibegyam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitzkaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Wellenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
DR EMBL: AE003450; AAF48196.1; -
DR FlyBase: FBgn0024308; Smr.
DR InterPro: IPR001004; -
DR InterPro: IPR001005; -
DR InterPro: IPR002086; -
DR Pfam: PF00249; myb_DNA-binding; 1.
DR PRINTS: PR00308; ANTIREFEZET.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; 1.
DR SMART: SM00395; SANT; 1.
SQ SEQUENCE 3502 AA; 369068 MW; 74C8004F9DA8FB9D CRC64;

Query Match 54.1%; Score 53; DB 5; Length 3502;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HGHGCHKHKN 11
DB 1929 HGHGCHKHKN 1938

RESULT 12

O9M435 PRELIMINARY; PRT; 79 AA.
AC O9M435;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PHASE-CHANGE RELATED PROTEIN PRECURSOR.

DR FlyBase; FBgn0003053; pzb.
DR InterPro; IPRO00822; -
DR Pfam; PF00096; zf-C2H2; 12.
DR PRINTS; PR00048; ZINC_FINGER.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 11.
DR SMART; SM00355; ZNF_C2H2; 1.
DR DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ SEQUENCE 1891 AA; 205368 MW; 9E882364C36BB9DF CRC64;

Query Match 55.1%; Score 54; DB 5; Length 1891;
Best Local Similarity 66.7%; Pred. No. 5;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 2 HGHGSHGSHG 13
Db 1133 HGHGSHGSHG 1144

RESULT 8
O9W4J1 PRELIMINARY; PRT; 1893 AA.
AC O9W4J1; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PEB PROTEIN.
GN PEB OR EG:66A1.1 OR CG12212.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borzova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
RL EMBL; AE003431; AAF45960.1; -.
DR HSSP; P08046; 1A11.
DR FlyBase; FBgn0003053; pzb.
DR InterPro; IPRO00822; -
DR Pfam; PF00096; zf-C2H2; 12.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 11.
DR SMART; SM00355; ZNF_C2H2; 1.
DR DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 1893 AA; 205673 MW; 1B223EE96468A754 CRC64;

Query Match 55.1%; Score 54; DB 5; Length 1893;
Best Local Similarity 66.7%; Pred. No. 5;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 2 HGHGSHGSHG 13
Db 1136 HGHGSHGSHG 1147

RESULT 9
O9W2E8 PRELIMINARY; PRT; 110 AA.
AC O9W2E8; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG10320 PROTEIN.
GN CG10320.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borzova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RESULT 4
 ID Q27920 PRELIMINARY; PRT; 450 AA.
 AC Q27920;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, last annotation update)
 DE PC4-1.
 GN PC4-1.
 OS Bradysia hygida.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Scleroidea;
 OC Scleridae; Bradysia.
 OX NCBI_TaxID=35572;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SALIVARY GLAND.
 RX MEDLINE=95393845; PubMed=7664619;
 RA Moneci N., Fernandez M.A., Fontes A.M., Basso L.R., Nakanishi Y.,
 RA Baron B., Butlin G., Pao-Larson M.L.;
 RT "Molecular characterization of an 18 kb segment of DNA puf C4 of
 RT Bradysia hygida (Diptera: scleridae).";
 RL Chromosoma 103:715-724(1995).
 DR EMBL: U13883; AAA83554.1; -;
 DR EMBL: U13892; AAA83555.1; -;
 SQ SEQUENCE 450 AA; 47185 MW; 1F0633CE9B7F964C CRC64;

Query Match 56.1%; Score 55; DB 5; Length 450;
 Best Local Similarity 53.3%; Pred. No. 0.88;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 HGHGKHKHKKGN 16
 DB 87 HGHGKHNHGGKKN 101

RESULT 5
 ID Q9P6B0 PRELIMINARY; PRT; 125 AA.
 AC Q9P6B0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
 DE HYPOTHETICAL 13.3 KDA PROTEIN.
 GN BID1.100.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Allyn V., Hohnes J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL355927; CAB91259.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 125 AA; 13315 MW; BB0879A8491FDBD7 CRC64;

Query Match 55.6%; Score 54.5; DB 3; Length 125;
 Best Local Similarity 73.3%; Pred. No. 0.3;
 Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 2 HGHGKHKHKKGN 15
 DB 74 HDHGRKHKHKKGN 88

RESULT 6
 ID Q9U229 PRELIMINARY; PRT; 643 AA.
 AC Q9U229;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 DE Y56A3A.32 PROTEIN.
 GN Y56A3A.32.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."; Science 282:2012-2018(1998).
 RL -1- COFACTOR: FAD (BY SIMILARITY).
 CC -1- SIMILARITY: TO PYRIDINE NUCLEOTIDE-DISULPHIDE OXIDOREDUCTASES
 CC CLASS-1.
 DR EMBL: AL132860; CAB60511.1; -;
 DR InterPro: IPR001100; -;
 DR InterPro: IPR001327; -;
 DR Pfam: PF00070; PYR_redox; 1.
 DR PRINTS: PR00368; FADPNR.
 DR PRINTS: PR00411; PNDRTASE1.
 KW FAD; Flavoprotein; Oxidoreductase; Redox-active center.
 SQ SEQUENCE 643 AA; 71257 MW; EAC9A34980A5F75D CRC64;

Query Match 55.1%; Score 54; DB 5; Length 643;
 Best Local Similarity 72.7%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HGHGKHKHKKGN 12
 DB 106 HGHGKHKHKKGN 116

RESULT 7
 ID Q77275 PRELIMINARY; PRT; 1891 AA.
 AC Q77275;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
 DE EG:66A1.1 PROTEIN.
 GN PEB OR EG:66A1.1 OR CG12212.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ferraz C., Vidal S., Brun C., Bucheton A., Demallie J.G.;
 RT "Sequencing the distal x chromosome of Drosophila melanogaster."; Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Benos P.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 DR EMBL: AL031227; CAA20227.1; -;
 DR HSP; P08046; IAL1.

```

AC 09VK49;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CG16800 PROTEIN.
GN CG16800.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abrill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtils K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003638; AAF53230.1;
DR Flybase: FBgn0032462; CG16800.
SQ SEQUENCE 255 AA; 29296 MW; A52DF0CCDDF414DC CRC64;

```

```

Query Match 57.1%; Score 56; DB 5; Length 255;
Best Local Similarity 60.0%; Pred. No. 0.36;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 1 KHGHGKHKKNGK 15
DB 64 KHGHGKHKKNGK 78

```

```

RESULT 3
Q9VMS0 PRELIMINARY; PRT; 686 AA.
AC Q9VMS0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

```

```

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CG6632 PROTEIN.
GN CG6632.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abrill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtils K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003509; AAF48868.1;
DR Flybase: FBgn0030945; CG6632.
DR InterPro: IPR000104;
DR InterPro: IPR000169;
DR InterPro: IPR001965;
DR InterPro: IPR002395;
DR Pfam: PF00628; PHD; 1.
DR PRINTS: PR00308; ANTIPEEZEL.
DR PRINTS: PR00334; KININOGEN.
DR PROSITE: PS00639; THIOLEPROTEIN_HIS; UNKNOWN.1.
DR SMART: SM00249; PHD; 1.
SQ SEQUENCE 686 AA; 70647 MW; 17C56F19B5DD2901 CRC64;

```

```

Query Match 57.1%; Score 56; DB 5; Length 686;
Best Local Similarity 66.7%; Pred. No. 0.93;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 2 HGHGKHKKNG 13
DB 476 HGHGKHKKNG 487

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:25:55 ; Search time 118.42 Seconds

(without alignments)
17.876 Million cell updates/sec

Title: US-09-437-912-8

Perfect score: 98

Sequence: 1 KHGHCHGKHKNNKKN 16

Scoring table: BLOSUM62

Searched: 425026 segs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
1: SP:REMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_mus:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_unclassified:*
14: sp_vertebrate:*
15: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	66.3	126	11	009016
2	56	57.1	255	5	09VK49
3	55	57.1	686	5	09VWS0
4	55	56.1	450	5	027920
5	54.5	55.6	125	3	09PEB0
6	54	55.1	643	5	09U229
7	54	55.1	1891	5	077275
8	54	55.1	1893	5	09W4J1
9	53	54.1	110	5	09W2E8
10	53	54.1	3469	5	09U4I2
11	53	54.1	3502	5	09VYJ9
12	52	53.1	79	10	09M435
13	52	53.1	333	10	09M430
14	51.5	52.6	819	5	09VHR3
15	51	52.0	2230	5	09VY40
16	50.5	51.5	1493	5	09VEF7
17	50	51.0	229	5	09XU12
18	50	51.0	241	5	09XW21
19	50	51.0	241	5	09XW21

20	50	51.0	444	5	076472	076472 musca domes
21	50	51.0	550	5	017145	017145 lucilia cup
22	50	51.0	601	5	P92138	P92138 drosophila
23	50	51.0	606	5	09VSV0	09VSV0 drosophila
24	50	51.0	1018	5	09VX18	09VX18 drosophila
25	50	51.0	1235	5	09VYF3	09VYF3 drosophila
26	50	51.0	1920	5	046205	046205 drosophila
27	49.5	50.5	140	5	09XWP9	09XWP9 caenorhabd
28	49	50.0	126	10	09M4H2	09M4H2 vitis vinif
29	49	50.0	325	5	09VYM9	09VYM9 drosophila
30	49	50.0	563	5	09VPS6	09VPS6 drosophila
31	48.5	49.5	197	10	09SM38	09SM38 sporobolus
32	48.5	49.5	199	10	09LYB2	09LYB2 arabidopsis
33	48	49.0	110	10	064396	064396 plasm saliv
34	48	49.0	198	5	09NNV9	09NNV9 plasmodium
35	48	49.0	207	10	043397	043397 brassica na
36	48	49.0	251	5	018577	018577 caenorhabd
37	48	49.0	375	10	P93066	P93066 brassica na
38	48	49.0	484	2	09R7U0	09R7U0 escherichia
39	48	49.0	963	5	09U5X1	09U5X1 drosophila
40	48	49.0	1029	5	09N9H6	09N9H6 drosophila
41	48	49.0	1064	5	09VSN1	09VSN1 drosophila
42	48	49.0	1085	5	024455	024455 drosophila
43	48	49.0	1268	5	09V4I9	09V4I9 drosophila
44	48	49.0	1299	5	09U5X0	09U5X0 drosophila
45	47.5	48.5	142	5	09W152	09W152 drosophila

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	126 AA.
009016	009016			
AC	009016			
DT	01-JUL-1997 (TREMBLrel. 04, Created)			
DT	01-JUL-1997 (TREMBLrel. 04, last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, last annotation update)			
DE	K-KINNOGEN (FRAGMENT).			
GN	KNGK.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=DONRTU;			
RX	MEDLINE=97468288; PubMed=9321484;			
RA	Harris E.L., Grigor M.R., Innes B.A., Harlap S.B., Kolke G.,			
RA	Jacob H.J.;			
RT	*Strain-specific deletions in exon 10 of rat K-kininogen and T1-			
RT	kininogen genes allow mapping of both genes to rat chromosome 11.;			
RL	Mamm. Genome 8:791-792(1997).			
DR	EMBL: AF003623; AAC09070.1;			
DR	InterPro: IPR002395;			
DR	PRINTS: PRO0334; KININOGEN.			
FT	NON_TER	1		
FT	NON_TER	126		
SQ	SEQUENCE	126 AA; 14092 MW; 9CCDF8751DA49C88 CRC64;		
Query Match				
Best Local Similarity 66.3%; Score 65; DB 11; Length 126;				
Matches 12; Conservative 1; Mismatches 3; Indels 4; Gaps 1;				
QY	1 KHGHGHC---KHNNKGRN 16			
Db	78 RHGHGHCGRKTKNDKNN 97			
RESULT	2			
QYK49	PRELIMINARY:	PRT:	255 AA.	

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Query Match	47.4%	Score 46.5:	DB 1:	Length 352:
Best Local Similarity	47.8%	Pred. No. 5.5:		
Matches 11:	Conservative 2:	Mismatches 3:	Indels 7:	Gaps 2:
QY	1 KHGHG-HGKHKNK-----GKKN	16		
	11111	11111		
Db	98 KHGHADHGAERKRVVPEAGKN	120		

```
Search completed: July 6, 2001, 09:26:40
Job time: 971 sec
```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: U17133; AAA79234.1; -
 DR InterPro: IPR002524; -
 DR Pfam: PF01545; Cation_efflux; 1.
 KW Zinc; Transport; Transmembrane; Multigene family; Repeat.
 FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 11 31 POTENTIAL.
 FT TRANSMEM 32 35 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 36 56 POTENTIAL.
 FT TRANSMEM 57 78 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 79 99 POTENTIAL.
 FT TRANSMEM 100 113 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 135 247 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 248 268 POTENTIAL.
 FT TRANSMEM 269 307 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 308 328 POTENTIAL.
 FT TRANSMEM 329 507 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 145 156 6 X 2 AA APPROXIMATE REPEATS OF H-G.
 FT TRANSMEM 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 298 298
 SQ SEQUENCE 507 AA; 55142 MW; 9F9770017C2455FC CRC64;

Query Match 48.5%; Score 47.5; DB 1; Length 507;
 Best Local Similarity 64.3%; Pred. No. 5.6;
 Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 2 HGHGCHKHKKKK 15
 DB 150 HGHGCHKHKKKK 162

RESULT 14
 NR2M_HUMAN STANDARD; PRT; 97 AA.
 AC 043676;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE B12 SUBUNIT (EC 1.6.5.3) (EC 1.6.99.3)
 DE (COMPLEX I-B12) (CI-B12).
 GN NDUFB3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98086396; PubMed=9425316;
 RA Ton C., Hwang D.W., Dempsey A.A., Liew C.-C.;
 RT "Identification and primary structure of five human NADH-ubiquinone
 RT oxidoreductase subunits.";
 RL Biochem. Biophys. Res. Commun. 241:589-594(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Loeffen J., Smeets R., van den Heuvel L., Ruitenbeek W., Sengers R.,
 RA Trijbels F., Smeitink J.;
 RT "Human B12 subunit of NADH:ubiquinone oxidoreductase: cDNA cloning,
 RT chromosomal localization, tissue distribution, mutation detection, and
 RT the characterization of a B12 subunit pseudogene.";
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 CC TO BE UBIQUINONE.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.

CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF047183; AAC04268.1; -
 DR EMBL: AF035839; AAC15590.1; -
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Acetylation.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 97 AA; 11271 MW; 851709573DD5E586 CRC64;

Query Match 48.0%; Score 47; DB 1; Length 97;
 Best Local Similarity 70.0%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KHGHGCHKH 10
 DB 3 KHGHGCHKH 12

RESULT 15
 KE4_BRARE STANDARD; PRT; 352 AA.
 AC 09PGB8;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HISTIDINE-RICH MEMBRANE PROTEIN KE4 HOMOLOG (FRAGMENT).
 GN HK4.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Murray B.W., Smetlmann H., Klein J.;
 RT "Identification of a homolog of the human HK4 gene in zebrafish.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF196345; AAF05821.1; -
 KW Transmembrane; Glycoprotein.
 FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 128 148 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 348 368 POTENTIAL.
 FT TRANSMEM 388 408 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT TRANSMEM 468 488 POTENTIAL.
 FT TRANSMEM 508 528 POTENTIAL.
 FT TRANSMEM 548 568 POTENTIAL.
 FT TRANSMEM 588 608 POTENTIAL.
 FT TRANSMEM 628 648 POTENTIAL.
 FT TRANSMEM 668 688 POTENTIAL.
 FT TRANSMEM 708 728 POTENTIAL.
 FT TRANSMEM 748 768 POTENTIAL.
 FT TRANSMEM 788 808 POTENTIAL.
 FT TRANSMEM 828 848 POTENTIAL.
 FT TRANSMEM 868 888 POTENTIAL.
 FT TRANSMEM 908 928 POTENTIAL.
 FT TRANSMEM 948 968 POTENTIAL.
 FT TRANSMEM 988 1008 POTENTIAL.
 FT TRANSMEM 1028 1048 POTENTIAL.
 FT TRANSMEM 1068 1088 POTENTIAL.
 FT TRANSMEM 1108 1128 POTENTIAL.
 FT TRANSMEM 1148 1168 POTENTIAL.
 FT TRANSMEM 1188 1208 POTENTIAL.
 FT TRANSMEM 1228 1248 POTENTIAL.
 FT TRANSMEM 1268 1288 POTENTIAL.
 FT TRANSMEM 1308 1328 POTENTIAL.
 FT TRANSMEM 1348 1368 POTENTIAL.
 FT TRANSMEM 1388 1408 POTENTIAL.
 FT TRANSMEM 1428 1448 POTENTIAL.
 FT TRANSMEM 1468 1488 POTENTIAL.
 FT TRANSMEM 1508 1528 POTENTIAL.
 FT TRANSMEM 1548 1568 POTENTIAL.
 FT TRANSMEM 1588 1608 POTENTIAL.
 FT TRANSMEM 1628 1648 POTENTIAL.
 FT TRANSMEM 1668 1688 POTENTIAL.
 FT TRANSMEM 1708 1728 POTENTIAL.
 FT TRANSMEM 1748 1768 POTENTIAL.
 FT TRANSMEM 1788 1808 POTENTIAL.
 FT TRANSMEM 1828 1848 POTENTIAL.
 FT TRANSMEM 1868 1888 POTENTIAL.
 FT TRANSMEM 1908 1928 POTENTIAL.
 FT TRANSMEM 1948 1968 POTENTIAL.
 FT TRANSMEM 1988 2008 POTENTIAL.
 FT TRANSMEM 2028 2048 POTENTIAL.
 FT TRANSMEM 2068 2088 POTENTIAL.
 FT TRANSMEM 2108 2128 POTENTIAL.
 FT TRANSMEM 2148 2168 POTENTIAL.
 FT TRANSMEM 2188 2208 POTENTIAL.
 FT TRANSMEM 2228 2248 POTENTIAL.
 FT TRANSMEM 2268 2288 POTENTIAL.
 FT TRANSMEM 2308 2328 POTENTIAL.
 FT TRANSMEM 2348 2368 POTENTIAL.
 FT TRANSMEM 2388 2408 POTENTIAL.
 FT TRANSMEM 2428 2448 POTENTIAL.
 FT TRANSMEM 2468 2488 POTENTIAL.
 FT TRANSMEM 2508 2528 POTENTIAL.
 FT TRANSMEM 2548 2568 POTENTIAL.
 FT TRANSMEM 2588 2608 POTENTIAL.
 FT TRANSMEM 2628 2648 POTENTIAL.
 FT TRANSMEM 2668 2688 POTENTIAL.
 FT TRANSMEM 2708 2728 POTENTIAL.
 FT TRANSMEM 2748 2768 POTENTIAL.
 FT TRANSMEM 2788 2808 POTENTIAL.
 FT TRANSMEM 2828 2848 POTENTIAL.
 FT TRANSMEM 2868 2888 POTENTIAL.
 FT TRANSMEM 2908 2928 POTENTIAL.
 FT TRANSMEM 2948 2968 POTENTIAL.
 FT TRANSMEM 2988 3008 POTENTIAL.
 FT TRANSMEM 3028 3048 POTENTIAL.
 FT TRANSMEM 3068 3088 POTENTIAL.
 FT TRANSMEM 3108 3128 POTENTIAL.
 FT TRANSMEM 3148 3168 POTENTIAL.
 FT TRANSMEM 3188 3208 POTENTIAL.
 FT TRANSMEM 3228 3248 POTENTIAL.
 FT TRANSMEM 3268 3288 POTENTIAL.
 FT TRANSMEM 3308 3328 POTENTIAL.
 FT TRANSMEM 3348 3368 POTENTIAL.
 FT TRANSMEM 3388 3408 POTENTIAL.
 FT TRANSMEM 3428 3448 POTENTIAL.
 FT TRANSMEM 3468 3488 POTENTIAL.
 FT TRANSMEM 3508 3528 POTENTIAL.
 FT TRANSMEM 3548 3568 POTENTIAL.
 FT TRANSMEM 3588 3608 POTENTIAL.
 FT TRANSMEM 3628 3648 POTENTIAL.
 FT TRANSMEM 3668 3688 POTENTIAL.
 FT TRANSMEM 3708 3728 POTENTIAL.
 FT TRANSMEM 3748 3768 POTENTIAL.
 FT TRANSMEM 3788 3808 POTENTIAL.
 FT TRANSMEM 3828 3848 POTENTIAL.
 FT TRANSMEM 3868 3888 POTENTIAL.
 FT TRANSMEM 3908 3928 POTENTIAL.
 FT TRANSMEM 3948 3968 POTENTIAL.
 FT TRANSMEM 3988 4008 POTENTIAL.
 FT TRANSMEM 4028 4048 POTENTIAL.
 FT TRANSMEM 4068 4088 POTENTIAL.
 FT TRANSMEM 4108 4128 POTENTIAL.
 FT TRANSMEM 4148 4168 POTENTIAL.
 FT TRANSMEM 4188 4208 POTENTIAL.
 FT TRANSMEM 4228 4248 POTENTIAL.
 FT TRANSMEM 4268 4288 POTENTIAL.
 FT TRANSMEM 4308 4328 POTENTIAL.
 FT TRANSMEM 4348 4368 POTENTIAL.
 FT TRANSMEM 4388 4408 POTENTIAL.
 FT TRANSMEM 4428 4448 POTENTIAL.
 FT TRANSMEM 4468 4488 POTENTIAL.
 FT TRANSMEM 4508 4528 POTENTIAL.
 FT TRANSMEM 4548 4568 POTENTIAL.
 FT TRANSMEM 4588 4608 POTENTIAL.
 FT TRANSMEM 4628 4648 POTENTIAL.
 FT TRANSMEM 4668 4688 POTENTIAL.
 FT TRANSMEM 4708 4728 POTENTIAL.
 FT TRANSMEM 4748 4768 POTENTIAL.
 FT TRANSMEM 4788 4808 POTENTIAL.
 FT TRANSMEM 4828 4848 POTENTIAL.
 FT TRANSMEM 4868 4888 POTENTIAL.
 FT TRANSMEM 4908 4928 POTENTIAL.
 FT TRANSMEM 4948 4968 POTENTIAL.
 FT TRANSMEM 4988 5008 POTENTIAL.
 FT TRANSMEM 5028 5048 POTENTIAL.
 FT TRANSMEM 5068 5088 POTENTIAL.
 FT TRANSMEM 5108 5128 POTENTIAL.
 FT TRANSMEM 5148 5168 POTENTIAL.
 FT TRANSMEM 5188 5208 POTENTIAL.
 FT TRANSMEM 5228 5248 POTENTIAL.
 FT TRANSMEM 5268 5288 POTENTIAL.
 FT TRANSMEM 5308 5328 POTENTIAL.
 FT TRANSMEM 5348 5368 POTENTIAL.
 FT TRANSMEM 5388 5408 POTENTIAL.
 FT TRANSMEM 5428 5448 POTENTIAL.
 FT TRANSMEM 5468 5488 POTENTIAL.
 FT TRANSMEM 5508 5528 POTENTIAL.
 FT TRANSMEM 5548 5568 POTENTIAL.
 FT TRANSMEM 5588 5608 POTENTIAL.
 FT TRANSMEM 5628 5648 POTENTIAL.
 FT TRANSMEM 5668 5688 POTENTIAL.
 FT TRANSMEM 5708 5728 POTENTIAL.
 FT TRANSMEM 5748 5768 POTENTIAL.
 FT TRANSMEM 5788 5808 POTENTIAL.
 FT TRANSMEM 5828 5848 POTENTIAL.
 FT TRANSMEM 5868 5888 POTENTIAL.
 FT TRANSMEM 5908 5928 POTENTIAL.
 FT TRANSMEM 5948 5968 POTENTIAL.
 FT TRANSMEM 5988 6008 POTENTIAL.
 FT TRANSMEM 6028 6048 POTENTIAL.
 FT TRANSMEM 6068 6088 POTENTIAL.
 FT TRANSMEM 6108 6128 POTENTIAL.
 FT TRANSMEM 6148 6168 POTENTIAL.
 FT TRANSMEM 6188 6208 POTENTIAL.
 FT TRANSMEM 6228 6248 POTENTIAL.
 FT TRANSMEM 6268 6288 POTENTIAL.
 FT TRANSMEM 6308 6328 POTENTIAL.
 FT TRANSMEM 6348 6368 POTENTIAL.
 FT TRANSMEM 6388 6408 POTENTIAL.
 FT TRANSMEM 6428 6448 POTENTIAL.
 FT TRANSMEM 6468 6488 POTENTIAL.
 FT TRANSMEM 6508 6528 POTENTIAL.
 FT TRANSMEM 6548 6568 POTENTIAL.
 FT TRANSMEM 6588 6608 POTENTIAL.
 FT TRANSMEM 6628 6648 POTENTIAL.
 FT TRANSMEM 6668 6688 POTENTIAL.
 FT TRANSMEM 6708 6728 POTENTIAL.
 FT TRANSMEM 6748 6768 POTENTIAL.
 FT TRANSMEM 6788 6808 POTENTIAL.
 FT TRANSMEM 6828 6848 POTENTIAL.
 FT TRANSMEM 6868 6888 POTENTIAL.
 FT TRANSMEM 6908 6928 POTENTIAL.
 FT TRANSMEM 6948 6968 POTENTIAL.
 FT TRANSMEM 6988 7008 POTENTIAL.
 FT TRANSMEM 7028 7048 POTENTIAL.
 FT TRANSMEM 7068 7088 POTENTIAL.
 FT TRANSMEM 7108 7128 POTENTIAL.
 FT TRANSMEM 7148 7168 POTENTIAL.
 FT TRANSMEM 7188 7208 POTENTIAL.
 FT TRANSMEM 7228 7248 POTENTIAL.
 FT TRANSMEM 7268 7288 POTENTIAL.
 FT TRANSMEM 7308 7328 POTENTIAL.
 FT TRANSMEM 7348 7368 POTENTIAL.
 FT TRANSMEM 7388 7408 POTENTIAL.
 FT TRANSMEM 7428 7448 POTENTIAL.
 FT TRANSMEM 7468 7488 POTENTIAL.
 FT TRANSMEM 7508 7528 POTENTIAL.
 FT TRANSMEM 7548 7568 POTENTIAL.
 FT TRANSMEM 7588 7608 POTENTIAL.
 FT TRANSMEM 7628 7648 POTENTIAL.
 FT TRANSMEM 7668 7688 POTENTIAL.
 FT TRANSMEM 7708 7728 POTENTIAL.
 FT TRANSMEM 7748 7768 POTENTIAL.
 FT TRANSMEM 7788 7808 POTENTIAL.
 FT TRANSMEM 7828 7848 POTENTIAL.
 FT TRANSMEM 7868 7888 POTENTIAL.
 FT TRANSMEM 7908 7928 POTENTIAL.
 FT TRANSMEM 7948 7968 POTENTIAL.
 FT TRANSMEM 7988 8008 POTENTIAL.
 FT TRANSMEM 8028 8048 POTENTIAL.
 FT TRANSMEM 8068 8088 POTENTIAL.
 FT TRANSMEM 8108 8128 POTENTIAL.
 FT TRANSMEM 8148 8168 POTENTIAL.
 FT TRANSMEM 8188 8208 POTENTIAL.
 FT TRANSMEM 8228 8248 POTENTIAL.
 FT TRANSMEM 8268 8288 POTENTIAL.
 FT TRANSMEM 8308 8328 POTENTIAL.
 FT TRANSMEM 8348 8368 POTENTIAL.
 FT TRANSMEM 8388 8408 POTENTIAL.
 FT TRANSMEM 8428 8448 POTENTIAL.
 FT TRANSMEM 8468 8488 POTENTIAL.
 FT TRANSMEM 8508 8528 POTENTIAL.
 FT TRANSMEM 8548 8568 POTENTIAL.
 FT TRANSMEM 8588 8608 POTENTIAL.
 FT TRANSMEM 8628 8648 POTENTIAL.
 FT TRANSMEM 8668 8688 POTENTIAL.
 FT TRANSMEM 8708 8728 POTENTIAL.
 FT TRANSMEM 8748 8768 POTENTIAL.
 FT TRANSMEM 8788 8808 POTENTIAL.
 FT TRANSMEM 8828 8848 POTENTIAL.
 FT TRANSMEM 8868 8888 POTENTIAL.
 FT TRANSMEM 8908 8928 POTENTIAL.
 FT TRANSMEM 8948 8968 POTENTIAL.
 FT TRANSMEM 8988 9008 POTENTIAL.
 FT TRANSMEM 9028 9048 POTENTIAL.
 FT TRANSMEM 9068 9088 POTENTIAL.
 FT TRANSMEM 9108 9128 POTENTIAL.
 FT TRANSMEM 9148 9168 POTENTIAL.
 FT TRANSMEM 9188 9208 POTENTIAL.
 FT TRANSMEM 9228 9248 POTENTIAL.
 FT TRANSMEM 9268 9288 POTENTIAL.
 FT TRANSMEM 9308 9328 POTENTIAL.
 FT TRANSMEM 9348 9368 POTENTIAL.
 FT TRANSMEM 9388 9408 POTENTIAL.
 FT TRANSMEM 9428 9448 POTENTIAL.
 FT TRANSMEM 9468 9488 POTENTIAL.
 FT TRANSMEM 9508 9528 POTENTIAL.
 FT TRANSMEM 9548 9568 POTENTIAL.
 FT TRANSMEM 9588 9608 POTENTIAL.
 FT TRANSMEM 9628 9648 POTENTIAL.
 FT TRANSMEM 9668 9688 POTENTIAL.
 FT TRANSMEM 9708 9728 POTENTIAL.
 FT TRANSMEM 9748 9768 POTENTIAL.
 FT TRANSMEM 9788 9808 POTENTIAL.
 FT TRANSMEM 9828 9848 POTENTIAL.
 FT TRANSMEM 9868 9888 POTENTIAL.
 FT TRANSMEM 9908 9928 POTENTIAL.
 FT TRANSMEM 9948 9968 POTENTIAL.
 FT TRANSMEM 9988 10008 POTENTIAL.
 FT TRANSMEM 10028 10048 POTENTIAL.
 FT TRANSMEM 10068 10088 POTENTIAL.
 FT TRANSMEM 10108 10128 POTENTIAL.
 FT TRANSMEM 10148 10168 POTENTIAL.
 FT TRANSMEM 10188 10208 POTENTIAL.
 FT TRANSMEM 10228 10248 POTENTIAL.
 FT TRANSMEM 10268 10288 POTENTIAL.
 FT TRANSMEM 10308 10328 POTENTIAL.
 FT TRANSMEM 10348 10368 POTENTIAL.
 FT TRANSMEM 10388 10408 POTENTIAL.
 FT TRANSMEM 10428 10448 POTENTIAL.
 FT TRANSMEM 10468 10488 POTENTIAL.
 FT TRANSMEM 10508 10528 POTENTIAL.
 FT TRANSMEM 10548 10568 POTENTIAL.
 FT TRANSMEM 10588 10608 POTENTIAL.
 FT TRANSMEM 10628 10648 POTENTIAL.
 FT TRANSMEM 10668 10688 POTENTIAL.
 FT TRANSMEM 10708 10728 POTENTIAL.
 FT TRANSMEM 10748 10768 POTENTIAL.
 FT TRANSMEM 10788 10808 POTENTIAL.
 FT TRANSMEM 10828 10848 POTENTIAL.
 FT TRANSMEM 10868 10888 POTENTIAL.
 FT TRANSMEM 10908 10928 POTENTIAL.
 FT TRANSMEM 10948 10968 POTENTIAL.
 FT TRANSMEM 10988 11008 POTENTIAL.
 FT TRANSMEM 11028 11048 POTENTIAL.
 FT TRANSMEM 11068 11088 POTENTIAL.
 FT TRANSMEM 11108 11128 POTENTIAL.
 FT TRANSMEM 11148 11168 POTENTIAL.
 FT TRANSMEM 11188 11208 POTENTIAL.
 FT TRANSMEM 11228 11248 POTENTIAL.
 FT TRANSMEM 11268 11288 POTENTIAL.
 FT TRANSMEM 11308 11328 POTENTIAL.
 FT TRANSMEM 11348 11368 POTENTIAL.
 FT TRANSMEM 11388 11408 POTENTIAL.
 FT TRANSMEM 11428 11448 POTENTIAL.
 FT TRANSMEM 11468 11488 POTENTIAL.
 FT TRANSMEM 11508 11528 POTENTIAL.
 FT TRANSMEM 11548 11568 POTENTIAL.
 FT TRANSMEM 11588 11608 POTENTIAL.
 FT TRANSMEM 11628 11648 POTENTIAL.
 FT TRANSMEM 11668 11688 POTENTIAL.
 FT TRANSMEM 11708 11728 POTENTIAL.
 FT TRANSMEM 11748 11768 POTENTIAL.
 FT TRANSMEM 11788 11808 POTENTIAL.
 FT TRANSMEM 11828 11848 POTENTIAL.
 FT TRANSMEM 11868 11888 POTENTIAL.
 FT TRANSMEM 11908 11928 POTENTIAL.
 FT TRANSMEM 11948 11968 POTENTIAL.
 FT TRANSMEM 11988 12008 POTENTIAL.
 FT TRANSMEM 12028 12048 POTENTIAL.
 FT TRANSMEM 12068 12088 POTENTIAL.
 FT TRANSMEM 12108 12128 POTENTIAL.
 FT TRANSMEM 12148 12168 POTENTIAL.
 FT TRANSMEM 12188 12208 POTENTIAL.
 FT TRANSMEM 12228 12248 POTENTIAL.
 FT TRANSMEM 12268 12288 POTENTIAL.
 FT TRANSMEM 12308 12328 POTENTIAL.
 FT TRANSMEM 12348 12368 POTENTIAL.
 FT TRANSMEM 12388 12408 POTENTIAL.
 FT TRANSMEM 12428 12448 POTENTIAL.
 FT TRANSMEM 12468 12488 POTENTIAL.
 FT TRANSMEM 12508 12528 POTENTIAL.
 FT TRANSMEM 12548 12568 POTENTIAL.
 FT TRANSMEM 12588 12608 POTENTIAL.
 FT TRANSMEM 12628 12648 POTENTIAL.
 FT TRANSMEM 12668 12688 POTENTIAL.
 FT TRANSMEM 12708 12728 POTENTIAL.
 FT TRANSMEM 12748 12768 POTENTIAL.
 FT TRANSMEM 12788 12808 POTENTIAL.
 FT TRANSMEM 12828 12848 POTENTIAL.
 FT TRANSMEM 12868 12888 POTENTIAL.
 FT TRANSMEM 12908 12928 POTENTIAL.
 FT TRANSMEM 12948 12968 POTENTIAL.
 FT TRANSMEM 12988 13008 POTENTIAL.
 FT TRANSMEM 13028 13048 POTENTIAL.
 FT TRANSMEM 13068 13088 POTENTIAL.
 FT TRANSMEM 13108 13128 POTENTIAL.
 FT TRANSMEM 13148 13168 POTENTIAL.
 FT TRANSMEM 13188 13208 POTENTIAL.
 FT TRANSMEM 13228 13248 POTENTIAL.
 FT TRANSMEM 13268 13288 POTENTIAL.
 FT TRANSMEM 13308 13328 POTENTIAL.
 FT TRANSMEM 13348 13368 POTENTIAL.
 FT TRANSMEM 13388 13408 POTENTIAL.
 FT TRANSMEM 13428 13448 POTENTIAL.
 FT TRANSMEM 13468 13488 POTENTIAL.
 FT TRANSMEM 13508 13528 POTENTIAL.
 FT TRANSMEM 13548 13568 POTENTIAL.
 FT TRANSMEM 13588 13608 POTENTIAL.
 FT TRANSMEM 13628 13648 POTENTIAL.
 FT TRANSMEM 13668 13688 POTENTIAL.
 FT TRANSMEM 13708 13728 POTENTIAL.
 FT TRANSMEM 13748 13768 POTENTIAL.
 FT TRANSMEM 13788 13808 POTENTIAL.
 FT TRANSMEM 13828 13848 POTENTIAL.
 FT TRANSMEM 13868 13888 POTENTIAL.
 FT TRANSMEM 13908 13928 POTENTIAL.
 FT TRANSMEM 13948 13968 POTENTIAL.
 FT TRANSMEM 13988 14008 POTENTIAL.
 FT TRANSMEM 14028 14048 POTENTIAL.
 FT TRANSMEM 14068 14088 POTENTIAL.
 FT TRANSMEM 14108 14128 POTENTIAL.
 FT TRANSMEM 14148 14168 POTENTIAL.
 FT TRANSMEM 14188 14208 POTENTIAL.
 FT TRANSMEM 14228 14248 POTENTIAL.
 FT TRANSMEM 14268 14288 POTENTIAL.
 FT TRANSMEM 14308 14328 POTENTIAL.
 FT TRANSMEM 14348 14368 POTENTIAL.
 FT TRANSMEM 14388 14408 POTENTIAL.
 FT TRANSMEM 14428 14448 POTENTIAL.
 FT TRANSMEM 14468 14488 POTENTIAL.
 FT TRANSMEM 14508 14528 POTENTIAL.
 FT TRANSMEM 14548 14568 POTENTIAL.
 FT TRANSMEM 14588 14608 POTENTIAL.
 FT TRANSMEM 14628 14648 POTENTIAL.
 FT TRANSMEM 14668 14688 POTENTIAL.
 FT TRANSMEM 14708 14728 POTENTIAL.
 FT TRANSMEM 14748 14768 POTENTIAL.
 FT TRANSMEM 14788 14808 POTENTIAL.
 FT TRANSMEM 14828 14848 POTENTIAL.
 FT TRANSMEM 14868 14888 POTENTIAL.
 FT TRANSMEM 14908 14928 POTENTIAL.
 FT TRANSMEM 14948 14968 POTENTIAL.
 FT TRANSMEM 14988 15008 POTENTIAL.
 FT TRANSMEM 15028 15048 POTENTIAL.
 FT TRANSMEM 15068 15088 POTENTIAL.
 FT TRANSMEM 15108 15128 POTENTIAL.
 FT TRANSMEM 15148 15168 POTENTIAL.
 FT TRANSMEM 15188 15208 POTENTIAL.
 FT TRANSMEM 15228 15248 POTENTIAL.
 FT TRANSMEM 15268 15288 POTENTIAL.
 FT TRANSMEM 15308 15328 POTENTIAL.
 FT TRANSMEM 15348 15368 POTENTIAL.
 FT TRANSMEM 15388 15408 POTENTIAL.
 FT TRANSMEM 15428 15448 POTENTIAL.
 FT TRANSMEM 15468 15488 POTENTIAL.
 FT TRANSMEM 15508 15528 POTENTIAL.
 FT TRANSMEM 15548 15568 POTENTIAL.
 FT TRANSMEM 15588 15608 POTENTIAL.
 FT TRANSMEM 15628 15648 POTENTIAL.
 FT TRANSMEM 15668 15688 POTENTIAL.
 FT TRANSMEM 15708 15728 POTENTIAL.
 FT TRANSMEM 15748 15768 POTENTIAL.
 FT TRANSMEM 15788 15808 POTENTIAL.
 FT TRANSMEM 15828 15848 POTENTIAL.
 FT TRANSMEM 15868 15888 POTENTIAL.
 FT TRANSMEM 15908 15928 POTENTIAL.
 FT TRANSMEM 15948 15968 POTENTIAL.
 FT TRANSMEM 15988 16008 POTENTIAL.
 FT TRANSMEM 16028 16048 POTENTIAL.
 FT TRANSMEM 16068 16088 POTENTIAL.
 FT TRANSMEM 16108 16128 POTENTIAL.
 FT TRANSMEM 16148 16168 POTENTIAL.
 FT TRANSMEM 16188 16208 POTENTIAL.
 FT TRANSMEM 16228 16248 POTENTIAL.
 FT TRANSMEM 16268 16288 POTENTIAL.
 FT TRANSMEM 16308 16328 POTENTIAL.
 FT TRANSMEM 16348 16368 POTENTIAL.
 FT TRANSMEM 16388 16408 POTENTIAL.
 FT TRANSMEM 16428 16448 POTENTIAL.
 FT TRANSMEM 16468 16488 POTENTIAL.
 FT TRANSMEM 16508 16528 POTENTIAL.
 FT TRANSMEM 16548 16568 POTENTIAL.
 FT TRANSMEM 16588 16608 POTENTIAL.
 FT TRANSMEM 16628 16648 POTENTIAL.
 FT TRANSMEM 16668 16688 POTENTIAL.
 FT TRANSMEM 16708 16728 POTENTIAL.
 FT TRANSMEM 16748 16768 POTENTIAL.
 FT TRANSMEM 16788 16808 POTENTIAL.
 FT TRANSMEM 16828 16848 POTENTIAL.
 FT TRANSMEM 16868 16888 POTENTIAL.
 FT TRANSMEM 16908 16928 POTENTIAL.
 FT TRANSMEM 16948 16968 POTENTIAL.
 FT TRANSMEM 16988 17008 POTENTIAL.
 FT TRANSMEM 17028 17048 POTENTIAL.
 FT TRANSMEM 17068 17088 POTENTIAL.
 FT TRANSMEM 17108 17128 POTENTIAL.
 FT TRANSMEM 17148 17168 POTENTIAL.
 FT TRANSMEM 17188 17208 POTENTIAL.
 FT TRANSMEM 17228 17248 POTENTIAL.
 FT TRANSMEM 17268 17288 POTENTIAL.
 FT TRANSMEM 17308 17328 POTENTIAL.
 FT TRANSMEM 17348 17368 POTENTIAL.
 FT TRANSMEM 17388 17408 POTENTIAL.
 FT TRANSMEM 17428 17448 POTENTIAL.
 FT TRANSMEM 17468 17488 POTENTIAL.
 FT TRANSMEM 17508 17528 POTENTIAL.
 FT TRANSMEM 17548 17568 POTENTIAL.
 FT TRANSMEM 17588 17608 POTENTIAL.
 FT TRANSMEM 17628 17648 POTENTIAL.
 FT TRANSMEM 17668 17688 POTENTIAL.
 FT TRANSMEM 17708 17728 POTENTIAL.
 FT TRANSMEM 17748 17768 POTENTIAL.
 FT TRANSMEM 17788 17808 POTENTIAL.
 FT TRANSMEM 17828 17848 POTENTIAL.
 FT TRANSMEM 17868 17888 POTENTIAL.
 FT TRANSMEM 17908 17928 POTENTIAL.
 FT TRANSMEM 17948 17968 POTENTIAL.
 FT TRANSMEM 17988 18008 POTENTIAL.
 FT TRANSMEM 18028 18048 POTENTIAL.
 FT TRANSMEM 18068 18088 POTENTIAL.
 FT TRANSMEM 18108 18128 POTENTIAL.
 FT TRANSMEM 18148 18168 POTENTIAL.
 FT TRANSMEM 18188 18208 POTENTIAL.
 FT TRANSMEM 18228 18248 POTENTIAL.
 FT TRANSMEM 18268 18288 POTENTIAL.
 FT TRANSMEM 18308 18328 POTENTIAL.
 FT TRANSMEM 18348 18368 POTENTIAL.
 FT TRANSMEM 18388 18408 POTENTIAL.
 FT TRANSMEM 18428 18448 POTENTIAL.
 FT TRANSMEM 18468 18488 POTENTIAL.
 FT TRANSMEM 18508 18528 POTENTIAL.
 FT TRANSMEM 18548 1

CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE000162; AAC73671.1; -
DR EMBL: U84598; AAB40768.1; -
DR Ecogene: EG13642; ybc2.
DR InterPro: IPR000410; -
DR InterPro: IPR000658; -
DR Pfam: PF00672; DUF5; 1.
DR Pfam: PF00512; signal; 1.
KW Hypothetical protein; Sensory transduction; Transferase; Kinase;
KW Transmembrane; Inner membrane; Phosphorylation.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 36 POTENTIAL.
FT DOMAIN 37 186 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 187 207 POTENTIAL.
FT DOMAIN 208 480 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 271 271 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 480 AA; 53738 MW; B3D78C0810241519 CRC64;

Query Match
Best Local Similarity 49.0%; Score 48; DB 1; Length 480;
Best Local Similarity 70.0%; Pred. No. 4.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HGHGKHKHN 11
| | | | | : | |
DB 138 HGHGMEHSN 147

RESULT 12
ZNT1_MOUSE STANDARD; PRT; 503 AA.
ID ZNT1_MOUSE
AC Q60738;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ZINC TRANSPORTER 1 (ZNT-1).
GN SLC30A1 OR ZNT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=9518868; PubMed=7882967;
RA Palmer R.D., Findley S.D.;
RT *Cloning and functional characterization of a mammalian zinc
RT transporter that confers resistance to zinc.*;
RL EMBL J. 14:639-649(1995).
CC -1- FUNCTION: MAY BE INVOLVED IN ZINC TRANSPORT OUT OF THE CELL.
CC LETHALITY OF KNOCKOUT EARLY IN GESTATION SUGGESTS A ROLE OF THE
CC PROTEIN IN FETAL ZINC ACQUISITION AND RETENTION.
CC -1- SUBUNIT: MULTIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC LOCALIZED ON THE PLASMA MEMBRANE (PROBABLE).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE SLC30A FAMILY OF TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U17132; AAA79233.1; -
DR MGD: MGI:1345281; SLC30a1.
DR InterPro: IPR002524; -
DR Pfam: PF01545; Cation_efflux; 1.
KW Zinc; Transport; Transmembrane; Multigene family; Repeat.
FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 11 31 POTENTIAL.
FT DOMAIN 32 35 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 36 56 POTENTIAL.
FT DOMAIN 57 78 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 79 99 POTENTIAL.
FT DOMAIN 100 113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 114 134 POTENTIAL.
FT DOMAIN 135 243 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 244 264 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 265 303 POTENTIAL.
FT TRANSMEM 304 324 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 325 503 POTENTIAL.
FT TRANSMEM 145 503 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 294 294 6 X 2 AA APPROXIMATE REPEATS OF H-G.
FT N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 503 AA; 54716 MW; 7C4FF93FC13CDA22 CRC64;

Query Match
Best Local Similarity 48.5%; Score 47.5; DB 1; Length 503;
Best Local Similarity 64.3%; Pred. No. 5.6;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 2 HGHGKHKHNKKR 15
| | | | | | | | | | | | | | | | | |
DB 150 HGHGHC-HLAKGAR 162

RESULT 13
ZNT1_RAT STANDARD; PRT; 507 AA.
ID ZNT1_RAT
AC Q62720;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ZINC TRANSPORTER 1 (ZNT-1).
GN SLC30A1 OR ZNT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=9518868; PubMed=7882967;
RA Palmer R.D., Findley S.D.;
RT *Cloning and functional characterization of a mammalian zinc
RT transporter that confers resistance to zinc.*;
RL EMBL J. 14:639-649(1995).
RN [2]
RP INDUCTION BY ZINC.
RC TISSUE=Intestine;
RA MEDLINE=98226729; PubMed=9560190;
RA McMahon R.J., Cousins R.J.;
RT *Regulation of the zinc transporter ZNT-1 by dietary zinc.*;
RL Proc. Natl. Acad. Sci. U.S.A. 95:4841-4846(1998).
CC -1- FUNCTION: MAY BE INVOLVED IN ZINC TRANSPORT OUT OF THE CELL.
CC -1- SUBUNIT: MULTIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC LOCALIZED ON THE PLASMA MEMBRANE (PROBABLE). LOCALIZED ON THE
CC BASOLATERAL SURFACE OF THE ENTEROCYTES.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE PROTEIN IS DETECTED IN
CC DUODENUM AND JEJUNUM BUT NOT IN ILEUM AND COLON.
CC -1- INDUCTION: SLIGHTLY BY ZINC IN THE INTESTINE, BUT NOT THE LIVER.
CC -1- SIMILARITY: BELONGS TO THE SLC30A FAMILY OF TRANSPORTERS.
CC -----

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:17 ; Search time 113.68 Seconds
(without alignments)
6.399 Million cell updates/sec

Title: US-09-437-912-3

Perfect score: 79

Sequence: 1 GKKHKGHGCHK 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_0601.*
2: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	12	AA1994	Human high molecu
2	79	100.0	28	AA1997	Human high molecu
3	79	100.0	47	AA1997	Human high molecu
4	79	100.0	62	AA1997	Human high molecu
5	79	100.0	63	AA1997	Human high molecu
6	79	100.0	83	AA1997	Human high molecu
7	79	100.0	94	AA1997	Human high molecu
8	79	100.0	131	AA1997	Human high molecu
9	79	100.0	179	AA1997	Human high molecu
10	79	100.0	186	AA1997	Human high molecu
11	79	100.0	255	AA1997	Human high molecu

12	72	91.1	41	16	AA1997
13	72	91.1	110	16	AA1997
14	68	86.1	11	21	AA1997
15	62	78.5	309	21	AA1997
16	62	78.5	309	21	AA1997
17	62	78.5	330	21	AA1997
18	62	78.5	330	21	AA1997
19	62	78.5	344	21	AA1997
20	62	78.5	344	21	AA1997
21	62	78.5	389	21	AA1997
22	62	78.5	389	21	AA1997
23	62	78.5	398	21	AA1997
24	62	78.5	398	21	AA1997
25	62	78.5	425	21	AA1997
26	62	78.5	453	21	AA1997
27	60	75.9	20	17	AA1997
28	59	74.7	69	16	AA1997
29	55	69.6	564	19	AA1997
30	53	67.1	177	21	AA1997
31	53	67.1	180	19	AA1997
32	53	67.1	831	22	AA1997
33	52.5	66.5	293	21	AA1997
34	52.5	66.5	307	21	AA1997
35	52.5	66.5	490	21	AA1997
36	52	65.8	16	21	AA1997
37	52	65.8	19	21	AA1997
38	51	64.6	1213	17	AA1997
39	51	64.6	1213	18	AA1997
40	50	63.3	345	19	AA1997
41	50	63.3	441	21	AA1997
42	50	63.3	473	21	AA1997
43	50	63.3	488	21	AA1997
44	49	62.0	20	17	AA1997
45	49	62.0	21	21	AA1997

ALIGNMENTS

RESULT 1
AA1994
ID AA1994 standard; peptide: 12 AA.
AC AA1994;
DT 16-OCT-2000 (first entry)
DE Human high molecular weight kininogen domain 5 fragment #3.
KW Human: high molecular weight kininogen; HK:
KW two-chain high molecular weight kininogen; HK:
KW angiotensin inhibition; tumour; cancer; ocular disorder;
KW rheumatoid arthritis; endothelial cell apoptosis.
OS Homo sapiens.
PN WO200027866-A1.
PP 18-MAY-2000.
PR 05-NOV-1999; 99WO-US26419.
PA 10-NOV-1998; 98US-0107833.
PE (UTEM) UNIV TEMPLE.
PI (MCCR) MCCR R K.
PI MCCR R K.
WP: 2000-376483/32.
A pharmaceutical composition used to inhibit angiogenesis, inhibit endothelial cell proliferation, and induce endothelial cell apoptosis

PT -
XX
PS Claim 10; Page 28; 52pp; English.
XX
CC The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain
CC high molecular weight kininogen (Hka) by plasma kallikrein. Hka or a
CC synthetic compound comprising part or all of the present sequence may
CC be used in a pharmaceutical composition for inhibiting angiogenesis.
CC Angiogenesis occurs in a number of disease states, such as tumour
CC formation and expansion, and certain ocular disorders. It can also
CC occur in a rheumatoid joint, hastening joint destruction by allowing
CC an influx of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial
CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.
XX
SQ Sequence 12 AA;
XX
Query Match 100.0%; Score 79; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GHKHKHGCHK 12
1 gnhkhghghgk 12
DB
RESULT 2
AAV81997
ID AAV81997 standard; peptide; 28 AA.
XX
XX AAV81997;
AC
XX
XX 16-OCT-2000 (first entry)
DT
XX
DE Human high molecular weight kininogen domain 5 fragment #6.
XX
XX Human; high molecular weight kininogen; HK;
KW two-chain high molecular weight kininogen; Hka;
KW angiogenesis inhibition; tumour; cancer; ocular disorder;
KW rheumatoid arthritis; endothelial cell apoptosis.
XX
XX Homo sapiens.
OS
XX
XX WO200027866-A1.
PN
XX
XX 18-MAY-2000.
PD
XX
XX 05-NOV-1999; 99WO-US26419.
PF
XX
XX 10-NOV-1998; 98US-0107833.
PR
XX
XX (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
PI
XX
XX McCrae RK;
PI
XX
XX WPI; 2000-376483/32.
DR
XX
XX A pharmaceutical composition used to inhibit angiogenesis, inhibit
PT endothelial cell proliferation, and induce endothelial cell apoptosis
PT
XX
XX -
XX
XX Claim 13; Page 29; 52pp; English.
XX
XX The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain

CC high molecular weight kininogen (Hka) by plasma kallikrein. Hka or a
CC synthetic compound comprising the present sequence may be used in a
CC pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
CC occurs in a number of disease states, such as tumour formation and
CC expansion, and certain ocular disorders. It can also occur in a
CC rheumatoid joint, hastening joint destruction by allowing an influx
CC of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial
CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.
XX
SQ Sequence 28 AA;
XX
Query Match 100.0%; Score 79; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GHKHKHGCHK 12
1 gnhkhghghgk 12
DB
RESULT 3
AAV93345
ID AAV93345 standard; peptide; 47 AA.
XX
XX AAV93345;
AC
XX
XX 04-SEP-2000 (first entry)
DT
XX
XX Light chain of human high molecular weight kininogen fragment.
DE
XX
XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
XX Synthetic.
OS
XX
XX Homo sapiens.
OS
XX
XX WO200027415-A2.
PN
XX
XX 18-MAY-2000.
PD
XX
XX 09-NOV-1999; 99WO-US26377.
PF
XX
XX 10-NOV-1998; 98US-0107844.
PR
XX
XX (UTEM) UNIV TEMPLE.
PA (DDPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MDS/) MDSA A S.
XX
XX Colman WR, Mousa AS;
PI
XX
XX WPI; 2000-376306/32.
DR
XX
XX Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
PT
XX
XX Claim 3; Page 36; 41pp; English.
XX
XX The present sequence represents a fragment of the light chain of human
CC high molecular weight kininogen. It is used to produce compounds of
CC the invention. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight

CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be
 CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.

XX Sequence 47 AA;

Query Match 100.0%; Score 79; DB 21; Length 47;

Best Local Similarity 100.0%; Pred. No. 9.2e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKKRKHGHGCK 12

DB 21 gkhkhghghgk 32

RESULT 4

AA93348

ID AAY93348 standard; peptide; 62 AA.

XX AAY93348;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

XX endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.

OS Homo sapiens.

PN WO200027415-A2.

PD 18-MAY-2000.

PF 09-NOV-1999; 99WO-US26377.

PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W. R.

PA (MOUS/) MOUSA A S.

PI Colman WR, Mousa AS;

DR WPI: 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
 PT that inhibit endothelial cell migration

XX Claim 6; Page 37; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a
 CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method
 CC of the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be
 CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.

XX Sequence 62 AA;

Query Match 100.0%; Score 79; DB 21; Length 62;

Best Local Similarity 100.0%; Pred. No. 1.2e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKKRKHGHGCK 12

DB 36 gkhkhghghgk 47

RESULT 5

AA975186

ID AAR75186 standard; peptide; 63 AA.

XX AAR75186;

DT 05-DEC-1995 (first entry)

DE Partial peptide of human HMW kininogen fragment 2.

KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 KW wound treating agent; bovine; growth promotion; fibroblast.

OS Homo sapiens.

PN JP07082172-A.

PD 28-MAR-1995.

PF 17-SEP-1993; 93JP-0230616.

PR 17-SEP-1993; 93JP-0230616.

PA (FARH) HOECHST JAPAN KK.

DR WPI: 1995-158909/21.

PT A wound treating agent contg. a partial peptide of kininogen -
 PT have growth promotion activity of fibroblasts.

PS Claim 8; Page 8; 8pp; Japanese.

CC AAR75186 is a partial peptide corresponding to human kininogen
 CC fragment 1, amino acids 458-520. Partial peptides of bovine and
 CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
 CC agent compsns. and act as the active component. The fragments are
 CC useful in wound treating because they have growth promotion activity
 CC on fibroblasts.

SQ Sequence 63 AA;

Query Match 100.0%; Score 79; DB 16; Length 63;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKKRKHGHGCK 12

DB 37 gkhkhghghgk 48

RESULT 6

AA93347

ID AAY93347 standard; peptide; 83 AA.

XX AAY93347;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

XX endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
PT WPI; 2000-376306/32.
DR
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
PS Claim 5; Page 37; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kda glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 83 AA;

Query Match 100.0%; Score 79; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHKHKHGHGK 12
Db 57 gnhkhghghgk 68

RESULT 7
AAV93351
ID AAV93351 standard; peptide; 94 AA.
XX
AC AAV93351;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; analogue; angiogenesis;
XX endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX

PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
PT WPI; 2000-376306/32.
DR
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
PS Claim 8; Page 39; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kda glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 94 AA;

Query Match 100.0%; Score 79; DB 21; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHKHKHGHGK 12
Db 57 gnhkhghghgk 68

RESULT 8
AAR75181
ID AAR75181 standard; peptide; 131 AA.
XX
AC AAR75181;
XX
DT 05-DEC-1995 (first entry)
XX
DE Partial peptide of human HMW kininogen fragment 1.2.
XX
KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KW wound treating agent; human; growth promotion; fibroblast.
XX
OS Homo sapiens.
XX
PN JP07082172-A.
XX
PD 28-MAR-1995.
XX
PF 17-SEP-1993; 93JP-0230616.
XX
PR 17-SEP-1993; 93JP-0230616.
XX
PA (FARH) HOECHST JAPAN KK.
XX
DR WPI; 1995-158909/21.
XX
PT A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
PS Claim 7; Page 7; 8pp; Japanese.
XX

CC AAR75181 is a partial peptide corresponding to human kininogen
CC fragment 1.2, amino acids 390-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compsns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.

XX
SQ Sequence 131 AA;

Query Match 100.0%; Score 79; DB 16; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKHKHGHGCK 12
|
Db 105 gnhkhghghgk 116

RESULT 9

AAV93353
ID AAV93353 standard; peptide; 179 AA.

XX
AC AAV93353;

XX
DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX
OS Synthetic.
OS Homo sapiens.

XX
PN WO200027415-A2.

XX
PD 18-MAY-2000.

XX
PF 09-NOV-1999; 99WO-US26377.

XX
PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.

XX
PI Colman WR, Mousa AS;

XX
DR WPI; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -

PS Claim 11; Page 40-41; 41pp; English.

XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cells proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.

XX
SQ Sequence 179 AA;

Query Match 100.0%; Score 79; DB 21; Length 179;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKHKHGHGCK 12
|
Db 29 gnhkhghghgk 40

RESULT 10

AAV93349
ID AAV93349 standard; peptide; 186 AA.

XX
AC AAV93349;

XX
DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX
OS Synthetic.
OS Homo sapiens.

XX
PN WO200027415-A2.

XX
PD 18-MAY-2000.

XX
PF 09-NOV-1999; 99WO-US26377.

XX
PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.

XX
PI Colman WR, Mousa AS;

XX
DR WPI; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -

PS Claim 9; Page 38; 41pp; English.

XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cells proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.

XX
SQ Sequence 186 AA;

Query Match 100.0%; Score 79; DB 21; Length 186;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKHKHGHGCK 12
|
Db 36 gnhkhghghgk 47

PA	(FARH) HOECHST JAPAN KK.
DR	WPI: 1995-158909/21.
XX	
XX	
PT	A wound treating agent confg. a partial peptide of kininogen -
PT	have growth promotion activity of fibroblasts.
XX	
PS	Claim 4; Page 6; 8pp; Japanese.
XX	
CC	AAR75178 is a partial peptide corresponding to bovine kininogen
CC	fragment 1.2, amino acids 387-496. Partial peptides of bovine and
CC	human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC	agent compsns. and act as the active component. The fragments are
CC	useful in wound treating because they have growth promotion activity
CC	on fibroblasts.
XX	
SO	Sequence 110 AA;
XX	
Query Match	91.1%; Score 72; DB 16; Length 110;
Best Local Similarity	91.7%; Pred. No. 0.00027;
Matches 11; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 GHKKHGHGCHK 12
Db	84 ghghkhghghgk 95
RESULT 14	
AA93352	
ID	AA93352 standard; peptide: 11 AA.
XX	
AC	AA93352;
XX	
DT	04-SEP-2000 (first entry)
XX	
DE	Light chain of human high molecular weight kininogen analogue.
XX	
KW	Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW	plasma kallikrein; heavy chain; light chain; analogue; angiotensin;
KW	endothelial cell proliferation; endothelial cell migration; vitronectin.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
PN	WO200027415-A2.
XX	
PD	18-MAY-2000.
XX	
PF	09-NOV-1999; 99WO-US26377.
XX	
PR	10-NOV-1998; 98US-0107844.
XX	
PA	(UTEM) UNIV TEMPLE.
PA	(DUPO) DUPONT PHARM CO.
PA	(COLM/) COLMAN W R.
PA	(MOUS/) MOUSA A S.
XX	
PI	Colman WR, Mousa AS;
XX	
DR	WPI: 2000-376306/32.
XX	
PT	Method for inhibiting endothelial cell proliferation, using compound
PT	that inhibit endothelial cell migration -
XX	
XX	
PS	Claim 10; Page 40; 41pp; English.
XX	
CC	The present sequence represents an analogue of the light chain of human
CC	high molecular weight kininogen. High molecular weight kininogen is a
CC	120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC	where it is cleaved by plasma kallikrein into heavy and light chains.
CC	Analogues of high molecular weight kininogen are used in the method
CC	of the invention. The specification describes a method of inhibiting

CC	endothelial cell proliferation. The method comprises contacting
CC	endothelial cells with a compound containing high molecular weight
CC	kinogen analogues. The method and the compounds can be used for
CC	inhibiting endothelial cell proliferation. The compounds can also be
CC	used for inhibiting angiogenesis. The compounds can also be used to
CC	inhibit migration of endothelial cells to vitronectin.
XX	
SQ	Sequence 11 AA;
Query Match	86.1%; Score 68; DB 21; Length 11;
Best Local Similarity	100.0%; Pred. No. 9.3e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GHKHKHGCH 10
Db	2 ghkhkhgh 11
RESULT 15	
AA606065	
ID	AA606065 standard; Protein; 309 AA.
XX	
AC	AA606065;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SPQ ID NO: 2701.
XX	
XX	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PM	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	28-APR-1999; 99US-0130891.
PR	30-APR-1999; 99US-0131444.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.

PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147491.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 21-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139817.	PR 30-AUG-1999;	99US-0151303.
PR 22-JUN-1999;	99US-0139899.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140353.	PR 01-SEP-1999;	99US-0151930.
PR 23-JUN-1999;	99US-0140354.	PR 07-SEP-1999;	99US-0152363.
PR 24-JUN-1999;	99US-0140695.	PR 10-SEP-1999;	99US-0153070.
PR 28-JUN-1999;	99US-0140823.	PR 13-SEP-1999;	99US-0153758.
PR 29-JUN-1999;	99US-0140991.	PR 15-SEP-1999;	99US-0154018.
PR 30-JUN-1999;	99US-0141287.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0141842.	PR 20-SEP-1999;	99US-0154779.
PR 01-JUL-1999;	99US-0142154.	PR 22-SEP-1999;	99US-0155139.
PR 02-JUL-1999;	99US-0142055.	PR 23-SEP-1999;	99US-0155486.
PR 06-JUL-1999;	99US-0142390.	PR 24-SEP-1999;	99US-0155659.
PR 08-JUL-1999;	99US-0142803.	PR 28-SEP-1999;	99US-0156458.
PR 09-JUL-1999;	99US-0142920.	PR 29-SEP-1999;	99US-0156596.
PR 12-JUL-1999;	99US-0142977.	PR 04-OCT-1999;	99US-0157117.
PR 13-JUL-1999;	99US-0143542.	PR 05-OCT-1999;	99US-0157753.
PR 14-JUL-1999;	99US-0143624.	PR 06-OCT-1999;	99US-0157865.
PR 15-JUL-1999;	99US-0144005.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144085.	PR 08-OCT-1999;	99US-0158232.
PR 16-JUL-1999;	99US-0144086.	PR 12-OCT-1999;	99US-0158368.
PR 19-JUL-1999;	99US-0144325.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144333.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159337.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161359.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161360.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146386.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.		

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.5%; Score 62; DB 21; Length 309;
Best Local Similarity 81.8%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKKKHGHG 11
1111111111
Db 125 ghahghghg 135

Search completed: July 6, 2001, 09:09:17
Job time: 123 sec

THIS PAGE BLANK (USPTO)

PT -
XX
PS Claim 14; Page 29; 52pp; English.
XX
CC The present sequence is derived from human two-chain high molecular
CC weight kininogen (Hka) domain 5. Hka is product of high molecular
CC weight kininogen (HK) cleavage by plasma kallikrein. HK is a 120 kD
CC glycoprotein which binds with high affinity to endothelial cells. Hka
CC or a synthetic compound comprising the present sequence may be
CC used in a pharmaceutical composition for inhibiting angiogenesis.
CC Angiogenesis occurs in a number of disease states, such as tumour
CC formation and expansion, and certain ocular disorders. It can also occur
CC in a rheumatoid joint, hastening joint destruction by allowing an influx
CC of leukocytes. The composition may inhibit angiogenesis by inhibiting
CC endothelial cell proliferation or by inducing endothelial cell
CC apoptosis. Peptides used in the composition may be recombinant peptides,
CC natural peptides, or synthetic peptides. They may also be chemically
CC synthesised, using, for example, solid phase synthesis methods.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 98; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHGHGKHKNGKKN 16
Db 1 khghgkhknkgkkn 16

RESULT 2

AAW07625
ID AAW07625 standard; peptide; 20 AA.

XX AC AAW07625;
XX DT 04-FEB-1997 (first entry)
XX

DE Human high polymer quininogen L-chain derived peptide.

KW Human; high polymer; quininogen; L-chain; cell adhesion;
KW cancer metastasis; platelet aggregation; inhibition; wound;
KW inflammatory disease; arteriosclerosis; glomerular nephritis;
KW treatment.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..13

FT Peptide /note= "claimed peptide (claim 1)"

FT Peptide 13..20

FT Peptide /note= "claimed peptide (claim 6)"

XX JP08208692-A.

XX PD 13-AUG-1996.

XX PF 28-SEP-1995; 95JP-0276418.

XX PR 28-SEP-1994; 94JP-0259451.

XX PA (SUMU) SUMITOMO SEIYAKU KK.

XX WIPI; 1996-421988/42.

XX DR Cell adhesion inhibiting peptide(s), used as cancer metastasis
XX PT inhibitor - comprises partial amino acid sequence of human high
XX PT polymer quininogen L chain

XX PS Claim 3; Page 2; 14pp; Japanese.

XX The present peptide, and its claimed fragments, are derived from

CC residues 402-498 of the human high polymer quininogen L-chain. They
CC are useful in cell adhesion, cancer metastasis or platelet
CC aggregation inhibitors, and in wound, inflammatory disease,
CC arteriosclerosis or glomerular nephritis treating agents. The
CC present peptide was synthesised using a solid phase method, and
CC purified using a YMC-DOS-120A-S15/13 column.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 98; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHGHGKHKNGKKN 16
Db 2 khghgkhknkgkkn 17

RESULT 3

AAV81997
ID AAV81997 standard; peptide; 28 AA.

XX AC AAV81997;
XX DT 16-OCT-2000 (first entry)

XX DE Human high molecular weight kininogen domain 5 fragment #6.

XX DE Human high molecular weight kininogen domain 5 fragment #6.

XX KW Human; high molecular weight kininogen; HK;
XX two-chain high molecular weight kininogen; Hka;
XX angiogenesis inhibition; tumour; cancer; ocular disorder;
XX rheumatoid arthritis; endothelial cell apoptosis.

XX KW Human; high molecular weight kininogen; HK;
XX two-chain high molecular weight kininogen; Hka;
XX angiogenesis inhibition; tumour; cancer; ocular disorder;
XX rheumatoid arthritis; endothelial cell apoptosis.

XX OS Homo sapiens.

XX PN WO200027866-A1.

XX PD 18-MAY-2000.

XX PF 05-NOV-1999; 99WO-US26419.

XX PR 10-NOV-1998; 98US-0107833.

XX PA (UTEM) UNIV TEMPLE.

XX PA (MCCR/) MCCR/ R K.

XX PI MCCR/ R K.

XX DR WIPI; 2000-376483/32.

XX PT A pharmaceutical composition used to inhibit angiogenesis, inhibit
XX PT endothelial cell proliferation, and induce endothelial cell apoptosis

XX PS Claim 13; Page 29; 52pp; English.

XX CC The present sequence is derived from human high molecular weight
XX CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
XX CC high affinity to endothelial cells, where it is cleaved to two-chain
XX CC high molecular weight kininogen (Hka) by plasma kallikrein. Hka or a
XX CC synthetic compound comprising the present sequence may be used in a
XX CC pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
XX CC occurs in a number of disease states, such as tumour formation and
XX CC expansion, and certain ocular disorders. It can also occur in a
XX CC rheumatoid joint, hastening joint destruction by allowing an influx
XX CC of leukocytes. The composition may inhibit angiogenesis by
XX CC inhibiting endothelial cell proliferation or by inducing endothelial
XX CC cell apoptosis. Peptides used in the composition may be recombinant
XX CC peptides, natural peptides, or synthetic peptides. They may also be
XX CC chemically synthesised, using, for example, solid phase synthesis
XX CC methods.

SQ Sequence 28 AA;

Query Match 100.0%; Score 98; DB 21; Length 28;

Best Local Similarity 100.0%; Pred. No. 2.8e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHGHGKHKNGKKN 16
|||||
Db 5 khghgkhkngkkn 20

RESULT 4

AA93345

ID AAY93345 standard; peptide; 47 AA.

XX AAY93345;

XX 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen fragment.

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;

KW Plasma kallikrein; heavy chain; analogue; angiogenesis;

KW Endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

OS Homo sapiens.

PN WO200027415-A2.

XX 18-MAY-2000.

PD 09-NOV-1999; 99WO-US26377.

XX 10-NOV-1998; 98US-0107844.

XX (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

XX Colman WR, Mousa AS;

XX MPI: 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
PI that inhibit endothelial cell migration

XX Claim 3; Page 36; 41pp; English.

CC The present sequence represents a fragment of the light chain of human
CC high molecular weight kininogen. It is used to produce compounds of
CC the invention. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells,

CC where it is cleaved by plasma kallikrein into heavy and light chains.

CC Analogues of high molecular weight kininogen are used in the method

CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting

CC endothelial cells with a compound containing high molecular weight

CC kininogen analogues. The method and the compounds can be used for

CC inhibiting endothelial cell proliferation. The compounds can also be

CC used for inhibiting angiogenesis. The compounds can also be used to

CC inhibit migration of endothelial cells to vitronectin.

XX Sequence 47 AA;

Query Match 100.0%; Score 98; DB 21; Length 47;

Best Local Similarity 100.0%; Pred. No. 4.7e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHGHGKHKNGKKN 16

Db |||||
25 khghgkhkngkkn 40

RESULT 5

AA93348

ID AAY93348 standard; peptide; 62 AA.

XX AAY93348;

XX 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;

KW Plasma kallikrein; heavy chain; analogue; angiogenesis;

KW Endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

OS Homo sapiens.

PN WO200027415-A2.

XX 18-MAY-2000.

PD 09-NOV-1999; 99WO-US26377.

XX 10-NOV-1998; 98US-0107844.

XX (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

XX Colman WR, Mousa AS;

XX MPI: 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
PI that inhibit endothelial cell migration

XX Claim 6; Page 37; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,

CC where it is cleaved by plasma kallikrein into heavy and light chains.

CC Analogues of high molecular weight kininogen are used in the method

CC of the invention. The specification describes a method of inhibiting

CC endothelial cell proliferation. The method comprises contacting

CC kininogen analogues. The method and the compounds can be used for

CC inhibiting endothelial cell proliferation. The compounds can also be

CC used for inhibiting angiogenesis. The compounds can also be used to

CC inhibit migration of endothelial cells to vitronectin.

XX Sequence 62 AA;

Query Match 100.0%; Score 98; DB 21; Length 62;

Best Local Similarity 100.0%; Pred. No. 6.3e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHGHGKHKNGKKN 16
|||||
Db 40 khghgkhkngkkn 55

RESULT 6

AA75186

ID AAR75186 standard; peptide; 63 AA.

XX AAR75186;

XX 05-DEC-1995 (first entry)
DT
XX
DE Partial peptide of human HMW kininogen fragment 2.
KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
XX wound treating agent; bovine; growth promotion; fibroblast.
OS Homo sapiens.
XX JP07082172-A.
PN
XX
XX 28-MAR-1995.
PD
XX 17-SEP-1993; 93JP-0230616.
PF
XX 17-SEP-1993; 93JP-0230616.
PR
XX (FARR) HOECHST JAPAN KK.
PA
XX WPI; 1995-158909/21.
DR
XX
PT A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
PS Claim 8; Page 8; 8pp; Japanese.
XX
CC AAR75186 is a partial peptide corresponding to human kininogen
CC fragment 1, amino acids 458-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compsns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 63 AA;

Query Match 100.0%; Score 98; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 6.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHGHGKHKKNKGKN 16
DB 41 khghgkhhkknkgkn 56

RESULT 7
AA93347
ID AAY93347 standard; peptide; 83 AA.
XX
AC AAY93347;
XX
XX 04-SEP-2000 (first entry)
DT
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin..
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA

PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
XX
XX WPI; 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
XX
PS Claim 5; Page 37; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 83 AA;

Query Match 100.0%; Score 98; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHGHGKHKKNKGKN 16
DB 61 khghgkhhkknkgkn 76

RESULT 8
AA93351
ID AAY93351 standard; peptide; 94 AA.
XX
AC AAY93351;
XX
XX 04-SEP-2000 (first entry)
DT
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin..
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
XX
XX WPI; 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -

XX PS Claim 8; Page 39; 41pp; English.
XX CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
SQ Sequence 94 AA;

Query Match 100.0%; Score 98; DB 21; Length 94;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHGHGCHKHKNKGKN 16
Db 61 khghgchkhknkgkn 76

RESULT 9
AAR75181
ID AAR75181 standard; peptide; 131 AA.
XX AC AAR75181;
XX DT 05-DEC-1995 (first entry)
XX DE Partial peptide of human HMW kininogen fragment 1.2.
XX DE high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
XX KW wound treating agent; human; growth promotion; fibroblast.
XX OS Homo sapiens.
XX PN JP07082172-A.
XX PD 28-MAR-1995.
XX PF 17-SEP-1993; 93JP-0230616.
XX PR 17-SEP-1993; 93JP-0230616.
XX PA (FARK) HOECHST JAPAN KK.
XX DR WPI; 1995-158909/21.
XX PT A wound treating agent contg. a partial peptide of kininogen -
XX PT have growth promotion activity of fibroblasts.
XX PS Claim 7; Page 7; 8pp; Japanese.
XX CC AAR75181 is a partial peptide corresponding to human kininogen
CC fragment 1.2, amino acids 390-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX SQ Sequence 131 AA;

QY 1 KHGHGCHKHKNKGKN 16
Db 109 khghgchkhknkgkn 124

RESULT 10
AAY93353
ID AAY93353 standard; peptide; 179 AA.
XX AC AAY93353;
XX DT 04-SEP-2000 (first entry)
XX DE Light chain of human high molecular weight kininogen analogue.
XX KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
XX KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
XX KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN MO200027415-A2.
XX PD 18-MAY-2000.
XX PF 09-NOV-1999; 99WO-US26377.
XX PR 10-NOV-1998; 98US-0107844.
XX PA (UTEM) UNIV TEMPLE.
XX PA (DUPO) DUPONT PHARM CO.
XX PA (COLM/) COLMAN W R.
XX PA (MOUS/) MOUSA A S.
XX PI Colman WR, Mousa AS;
XX DR WPI; 2000-376306/32.
XX PT Method for inhibiting endothelial cell proliferation, using compound
XX PT that inhibit endothelial cell migration -
XX PS Claim 11; Page 40-41; 41pp; English.
XX CC The present sequence represents an analogue of the light chain of human
XX CC high molecular weight kininogen. High molecular weight kininogen is a
XX CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
XX CC where it is cleaved by plasma kallikrein into heavy and light chains.
XX CC Analogues of high molecular weight kininogen are used in the method
XX CC of the invention. The specification describes a method of inhibiting
XX CC endothelial cell proliferation. The method comprises contacting
XX CC endothelial cells with a compound containing high molecular weight
XX CC kininogen analogues. The method and the compounds can be used for
XX CC inhibiting endothelial cell proliferation. The compounds can also be
XX CC used for inhibiting angiogenesis. The compounds can also be used to
XX CC inhibit migration of endothelial cells to vitronectin.
XX SQ Sequence 179 AA;

Query Match 100.0%; Score 98; DB 21; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHGHGCHKHKNKGKN 16
Db 33 khghgchkhknkgkn 48

RESULT 11
AAY93349
ID AAY93349 standard; peptide; 186 AA.

```

XX  AAY93349;
AC
XX
XX  04-SEP-2000 (first entry)
DT
XX
DE  Light chain of human high molecular weight kininogen analogue.
XX
XX  Human: high molecular weight kininogen; glycoprotein; endothelial cell;
KW  plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
XX  endothelial cell proliferation; endothelial cell migration; vitronectin.
OS
XX  Synthetic.
XX  Homo sapiens.
XX  WO200027415-A2.
XX
XX  18-MAY-2000.
XX
XX  09-NOV-1999; 99WO-US26377.
XX
XX  10-NOV-1998; 98US-0107844.
XX
XX  (UTEM ) UNIV TEMPLE.
PA  (DUPO ) DUPONT PHARM CO.
PA  (COLM/) COLMAN W R.
PA  (MOUS/) MOUSA A S.
XX
XX  Colman WR, Mousa AS;
PI
XX  WPI; 2000-376306/32.
DR
XX
XX  Method for inhibiting endothelial cell proliferation, using compound
PT  that inhibit endothelial cell migration
XX
XX  Claim 9; Page 38; 41pp; English.
PS
XX  The present sequence represents an analogue of the light chain of human
CC  high molecular weight kininogen. High molecular weight kininogen is a
CC  120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC  where it is cleaved by plasma kallikrein into heavy and light chains.
CC  Analogues of high molecular weight kininogen are used in the method
CC  of the invention. The specification describes a method of inhibiting
CC  endothelial cell proliferation. The method comprises contacting
CC  endothelial cells with a compound containing high molecular weight
CC  kininogen analogues. The method and the compounds can be used for
CC  inhibiting endothelial cell proliferation. The compounds can also be
CC  used for inhibiting angiogenesis. The compounds can also be used to
CC  inhibit migration of endothelial cells to vitronectin.
XX
XX  Sequence 186 AA;
SQ

```

```

Query Match 100.0%; Score 98; DB 21; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 KHGHGKHKNNKGN 16
    |||
DB 40 khghgkhknnkgn 55

```

```

RESULT 12
AAY93342
ID AAY93342 standard; protein; 255 AA.
XX
AC AAY93342;
XX
XX  04-SEP-2000 (first entry)
DT
XX
XX  Light chain of human high molecular weight kininogen.
DE
XX  Human: high molecular weight kininogen; glycoprotein; endothelial cell;
KW  plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
XX
XX

```

```

KW  endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
XX  Homo sapiens.
OS
XX  WO200027415-A2.
XX
XX  18-MAY-2000.
XX
XX  09-NOV-1999; 99WO-US26377.
XX
XX  10-NOV-1998; 98US-0107844.
XX
XX  (UTEM ) UNIV TEMPLE.
PA  (DUPO ) DUPONT PHARM CO.
PA  (COLM/) COLMAN W R.
PA  (MOUS/) MOUSA A S.
XX
XX  Colman WR, Mousa AS;
PI
XX  WPI; 2000-376306/32.
DR
XX
XX  Method for inhibiting endothelial cell proliferation, using compound
PT  that inhibit endothelial cell migration
XX
XX  Disclosure; Page 3; 41pp; English.
PS
XX  The present sequence represents the light chain of human high molecular
CC  weight kininogen. High molecular weight kininogen is a 120 kDa
CC  glycoprotein which binds with high affinity to endothelial cells,
CC  where it is cleaved by plasma kallikrein into heavy and light chains.
CC  Analogues of high molecular weight kininogen are used in the method
CC  of the invention. The specification describes a method of inhibiting
CC  endothelial cell proliferation. The method comprises contacting
CC  endothelial cells with a compound containing high molecular weight
CC  kininogen analogues. The method and the compounds can be used for
CC  inhibiting endothelial cell proliferation. The compounds can also be
CC  used for inhibiting angiogenesis. The compounds can also be used to
CC  inhibit migration of endothelial cells to vitronectin.
XX
XX  Sequence 255 AA;
SQ

```

```

Query Match 100.0%; Score 98; DB 21; Length 255;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 KHGHGKHKNNKGN 16
    |||
DB 109 khghgkhknnkgn 124

```

```

RESULT 13
AAR75180
ID AAR75180 standard; peptide; 41 AA.
XX
XX  AAR75180;
XX
XX  05-DEC-1995 (first entry)
DT
XX
XX  Partial peptide of HMW kininogen fragment 2.
DE
XX  high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KW  wound treating agent; bovine; growth promotion; fibroblast.
XX
XX  Bos taurus.
OS
XX  JP07082172-A.
XX
XX  28-MAR-1995.
XX
XX  17-SEP-1993; 93JP-0230616.
XX
XX  17-SEP-1993; 93JP-0230616.
XX

```

```

XX PA (FARH ) HOECHST JAPAN KK.
XX DR WPI; 1995-158909/21.
XX PT A wound treating agent contg. a partial peptide of kininogen -
XX PS have growth promotion activity of fibroblasts.
XX RS Claim 6; Page 7; 8pp; Japanese.
XX CC AAR75179 is a partial peptide corresponding to bovine kininogen
CC fragment 2, amino acids 456-496. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX SQ Sequence 41 AA;

Query Match 94.9%; Score 93; DB 16; Length 41;
Best Local Similarity 93.8%; Pred. No. 2e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KHGHGKHKNGKKN 16
   |||||
Db 19 khghgkhkngkgn 34

RESULT 14
AAR75178
ID AAR75178 standard; peptide; 110 AA.
XX AAR75178;
XX AC 05-DEC-1995 (first entry)
XX DT
XX DE Partial peptide of HMW kininogen fragment 1.2.
XX KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
XX KW wound treating agent; bovine; growth promotion; fibroblast.
XX OS Bos taurus.
XX FH Key Location/Qualifiers
XX FT Misc-difference 12 /label= Pro, Thr
XX FT Misc-difference 15 /label= Val or Leu
XX FT Misc-difference 69 /label= Lys or Arg
XX FT
XX PN JF07082172-A.
XX PD 28-MAR-1995.
XX PF 17-SEP-1993; 93JP-0230616.
XX PR 17-SEP-1993; 93JP-0230616.
XX PA (FARH ) HOECHST JAPAN KK.
XX DR WPI; 1995-158909/21.
XX CC A wound treating agent contg. a partial peptide of kininogen -
XX PT have growth promotion activity of fibroblasts.
XX PS Claim 4; Page 6; 8pp; Japanese.
XX CC AAR75178 is a partial peptide corresponding to bovine kininogen
CC fragment 1.2, amino acids 387-496. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compns. and act as the active component. The fragments are

```

CC	useful in wound treating because they have growth promotion activity
CC	on fibroblasts.
XX	
SQ	Sequence 110 AA;
Query Match	94.9%; Score 93; DB 16; Length 110;
Best Local Similarity	93.8%; Pred. NO. 5.4e-06;
Matches 15; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 KHGHGKKNKKGN 16
Db	88 khghgkknkkgn 103
RESULT 15	
AAY71879	
ID AAY71879 standard; peptide; 19 AA.	
XX	
AC AAY71879;	
XX	
DT 26-MAR-2001 (first entry)	
DE	
Human HKH20 peptide derived from domain 5 of H-Kinogen (479-498 aa).	
XX	
KW Human; heparin binding protein; HBP; antiinflammatory; cardiovascular;	
KW immunosuppressive; vasotropic; prevention; treatment; bradykinin;	
KW aprotinin; H-kininogen; HK; systemic inflammatory response syndrome;	
KW pre-kallikrein; ischaemia reperfusion; anaphylaxis; allograft rejection;	
KW adult respiratory distress syndrome; HKH20 peptide.	
XX	
OS Homo sapiens.	
PN	
WO200066151-A1.	
PD 09-NOV-2000.	
XX	
PF 28-APR-2000; 2000WO-DK00213.	
XX	
PR 29-APR-1999; 99US-0132748.	
PR 06-MAY-1999; 99DK-0000613.	
PR 01-OCT-1999; 99DK-0001402.	
PR 01-OCT-1999; 99US-0157384.	
XX	
PA (NOVO) NOVO NORDISK AS.	
XX	
PI Flodgaard HJ, Lindom L, Bjorn S;	
DR WPI; 2000-687445/67.	
XX	
Treating systemic inflammatory response syndrome, ischaemia reperfusion,	
anaphylaxis and allograft rejection by modulating release of bradykinin	
PT	
-	
Example 2, Page 39; 75pp; English.	
PS	
XX	
CC The patent discloses a method for preventing or treating a disorder	
CC resulting from the release of bradykinin in a mammal which produces	
CC a heparin-binding protein (HBP) that binds to a HBP antagonist. This	
CC method involves administration of a mammalian HBP antagonist (especially	
CC aprotinin) and/or monoclonal antibodies that bind to prekallikrein-	
CC H-kininogen complexes in the HBP, to decrease the release of bradykinin	
CC in the mammal. The antagonists of HBP (e.g. aprotinin) decrease the	
CC permeability of the endothelial cells and are used to prevent or treat	
CC disorders resulting from the release of bradykinin such as systemic	
CC inflammatory response syndrome, ischaemia reperfusion, anaphylaxis	
CC and/or allograft rejection. They are also used to treat adult	
CC respiratory distress syndrome.	
CC The present sequence is HKH20 peptide which is derived from the	
CC domain 5 of human H-kininogen (HK) protein (479-498 residues).	
CC HKH20 treatment of endothelial cells inhibits or prevents the HBP-	
XX induced increase in permeability of the endothelial cells.	

SQ Sequence 19 AA;

Query Match

Best Local Similarity 73.0%; Score 71.5; DB 21; Length 19;
 Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 KGHGSGKHKNGKKN 16
 ||||| |
 Db 2 khgphnk-kngkkn 16

Search completed: July 6, 2001, 09:09:19
 Job time: 125 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:19 ; Search time 113.68 Seconds
(without alignments)
8.533 Million cell updates/sec

Title: US-09-437-912-9

Perfect score: 98

Sequence: 1 HKNKGKNGKHNKMT 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601:*

- 1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
- 2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
- 3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
- 4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
- 5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
- 6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
- 7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
- 8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
- 9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
- 10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
- 11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
- 12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
- 13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
- 14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
- 15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
- 16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
- 17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
- 18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
- 19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
- 20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
- 21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
- 22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	16	21	Human two-chain hi
2	98	100.0	28	21	Human high molecu
3	98	100.0	94	21	Human high molecu
4	98	100.0	179	21	Human high molecu
5	98	100.0	186	21	Human high molecu
6	98	100.0	255	21	Human high molecu
7	93	94.9	47	21	Human high molecu
8	93	94.9	62	21	Human high molecu
9	93	94.9	63	16	Human high molecu
10	93	94.9	83	21	Human high molecu
11	93	94.9	131	16	Human high molecu

12	74	75.5	12	21	AAV81995
13	70	71.4	41	16	AAV75180
14	70	71.4	110	16	AAV75178
15	65	66.3	20	17	AAV07625
16	57	58.2	19	21	AAV71879
17	46	46.9	16	21	AAV81999
18	44	44.9	821	21	AAV42128
19	44	44.9	914	21	AAV42127
20	43	43.9	504	21	AAV81884
21	43	43.9	571	21	AAV30502
22	43	43.9	778	21	AAV30501
23	43	43.9	808	21	AAV30500
24	42	42.9	287	21	AAV38076
25	42	42.9	326	21	AAV40042
26	42	42.9	369	21	AAV38075
27	42	42.9	382	21	AAV03329
28	42	42.9	382	21	AAV40041
29	42	42.9	387	21	AAV38074
30	42	42.9	460	21	AAV05328
31	42	42.9	464	21	AAV40040
32	42	42.9	478	21	AAV05327
33	42	42.9	800	22	AAV71957
34	42	42.9	800	22	AAV65673
35	41	41.8	21	11	AAV04053
36	41	41.8	28	9	AAV81567
37	41	41.8	206	21	AAV07244
38	41	41.8	206	21	AAV50821
39	41	41.8	211	21	AAV31068
40	41	41.8	214	21	AAV07243
41	41	41.8	214	21	AAV50820
42	41	41.8	219	21	AAV31067
43	41	41.8	298	21	AAV07242
44	41	41.8	298	21	AAV50819
45	41	41.8	303	21	AAV31066

ALIGNMENTS

RESULT 1	
AAV06337	
ID	AAV06337 standard; Protein; 16 AA.
AC	AAV06337;
DT	16-OCT-2000 (first entry)
DE	Human two-chain high molecular weight kininogen domain 5 fragment #9.
KW	Human; high molecular weight kininogen; HK;
KW	two-chain high molecular weight kininogen; HKa;
KW	angiogenesis inhibition; tumour; cancer; ocular disorder;
KW	rheumatoid arthritis; endothelial cell apoptosis.
OS	Homo sapiens.
PN	WO200027866-A1*
PD	18-MAY-2000.
PR	10-NOV-1998; 98US-0107833.
PP	05-NOV-1999; 99WO-US26419.
PA	(UTEM) UNIV TEMPLE.
PI	(MCCR/) MCCRAE R K.
PI	MCCRae RK;
DR	WPI; 2000-376483/32.
PT	A pharmaceutical composition used to inhibit angiogenesis, inhibit endothelial cell proliferation, and induce endothelial cell apoptosis

Claim 15; Page 29; 52pp; English.

Claim 15, Page 29, 52pp: English.

50 Sequence 16 AA;

Query Match	100.0%;	Score 98;	DB 21;	Length 16;
Best Local Similarity	100.0%;	Pred. No. 2e-08;		
Matches 16;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	HKNKGGKNGKHNGWKT	16
Db	1	hknkgkngkhngwt	16

RESULT 2

ID AAY81997 standard; peptide; 28 AA.

AC AAY81997

DT 16-OCT-2000 (first entry)

Human high molecular weight kininogen domain 5 fragment #6

KM Human, high molecular weight kininogen; HK;
KM two-chain high molecular weight kininogen; Hka;
KM angiotensin inhibition; tumour; cancer; disorder
KM rheumatoid arthritis; endothelial cell apoptosis.

OS Homo sapiens.

PN WO2000027866-A1.

PD 18-MAY-2000.

PF 05-NOV-1999; 99WO-US26419

PR 10-NOV-1998; 98US-0107833

PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.

PI McCrae RK

DR WPI; 2000-376483/32.

PT A pharmaceutical composition used to inhibit angiogenesis, inhibit endothelial cell proliferation, and induce endothelial cell apoptosis

PS Claim 13; Page 29; 52pp; English.

CC The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain
CC high molecular weight kininogen (HKa) by plasma kallikrein. HKa or a

CC synthetic compound comprising the present sequence may be used in a
CC pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
CC occurs in a number of disease states, such as tumour formation and
CC expansion, and certain ocular disorders. It can also occur in a
CC rheumatoid joint, hastening joint destruction by allowing an influx
CC of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial
CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.

50 Sequence 28 Å.

Query Match	100.0%;	Score 98;	DB 21;	Length 28;
Best Local Similarity	100.0%;	Pred. No. 3.6e-08;		
Matches 16;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	HKNKGGKKNGKHNGWKT	16
Db	13	hkknkgkknghngwt	28

RESULT	3
AAY93351	

ID AAY93351

AC AAY93351;

DT	04-SEP-2000 (first entry)
XX	
DE	light chain of human high molecular weight kininogen analogue

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin

Synthetic.

XX
130300037415-23

XX 10 MAY 2000

XX 00-NOV-1 000. 00MO-JIC36377

XX 10-NOV-1 0000 0810-0107844

XX
XX
PA (ITEM) INTY TEMPI.F

PA	(DUPU) DUPONT PHA
PA	(COLM/) COLMAN W B

PA (MOUS/) MOUSA A S.
XX

PL COLLIDII WK, Mousd AS,
XX

PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -

PS Claim 8; Page 39; 41pp; English

CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be

CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.
 XX
 Sequence 94 AA;
 SQ

Query Match	100.0%;	Score 98;	DB 21;	Length 94;
Best Local Similarity	100.0%;	Pred. No. 1.2e-07;		
Matches 16;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	HNKNGKKNKGKHNKWT	16
Db	69	hknkgkknkgkhnkgwt	84

RESULT	4
AAV93353	
ID	AAV93353 standard; peptide; 179 AA.

AC	AAV93353;
XX	
DT	04-SEP-2000 (first entry)
XX	

Light chain of human high molecular weight kininogen analogue.

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell,
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin

OS Synthetic.
OS Homo sapiens.

PN W0200027415-A2.

PD 18-MAY-2000

PF 09-NOV-1999; 99WO-US26377

PR 10-NOV-1998; 98US-0107844

PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.

PA (MOUS/) MOUSA A S.

PI COLman WR, Mousa AS;

DR WPI; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -

PS Claim 11; Page 40-41; 41pp; English

The present sequence represents an analogue of the light chain of human high molecular weight kininogen. High molecular weight kininogen is a 120 kDa glycoprotein which binds with high affinity to endothelial cells where it is cleaved by plasma kallikrein into heavy and light chains. Analogues of high molecular weight kininogen are used in the method of the invention. The specification describes a method of inhibiting endothelial cell proliferation. The method comprises contacting endothelial cells with a compound containing high molecular weight kininogen analogues. The method and the compounds can be used for inhibiting endothelial cell proliferation. The compounds can also be used for inhibiting angiogenesis. The compounds can also be used to inhibit migration of endothelial cells to vitronectin.

SQ Sequence 179 AA;

Query Match	100.0%;	Score 98;	DB 21;	length 179;
Best Local Similarity	100.0%;	Pred. No. 2.4e-07;		
Matches 16; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

```
QY      1 HKNKGKKNGKNGWKT 16
         |||||
Db      41 hknkgkknghngwt 56
```

RESULT	5
AAV93349	
ID	AAV93349 standard; peptide; 186 AA.

AC	AA93349;
XX	
DT	04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

KM Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KM plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KM endothelial cell proliferation; endothelial cell migration; vitronectin

OS	Synthetic.
OS	Homo sapiens.

PN - W0200027415-A2

PD 18-MAY-2000

PF 09-NOV-1999; 99WO-US26377

PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.

PA (MOUS/) MOUSA A S.

PI Colman WR, Mousa AS,

DR WPT; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -

PS Claim 9; Page 38; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can be used to
CC inhibit migration of endothelial cells to vitronectin.

SQ Sequence 186 AA;

Query Match	100.0%	Score 98	DB 21	length 166
Best Local Similarity	100.0%	Pred. NC	2.5e-07	
Matches	16	Conservative	0	Mismatches 0; Indels 0; Gaps 0

OY	1	HKNKGKKNGKHNGWKT	16
Db	48	hknkgkknghngwt	63

RESULT	6
AAV93342	
ID	AAV93342 standard; protein; 255 AA

```

XX AC AAY93342;
XX XX
DT 04-SEP-2000 (first entry)
XX XX
DE Light chain of human high molecular weight kininogen.
XX XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX XX
OS Homo sapiens.
XX XX
PN WO200027415-A2.
XX XX
PD 18-MAY-2000.
XX XX
PF 09-NOV-1999; 99WO-US26377.
XX XX
PR 10-NOV-1998; 98US-0107844.
XX XX
PA (UTEM ) UNIV TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOSA/) MOUSA A S.
XX XX
PI Colman WR, Mousa AS;
XX XX
DR WPI; 2000-376306/32.
XX XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX XX
PS Disclosure; Page 3; 41pp; English.
XX XX
CC The present sequence represents the light chain of human high molecular
CC weight kininogen. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX XX
SQ Sequence 255 AA;

```

```

Query Match 100.0%; Score 98; DB 21; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HKNKGKNGKNGKNGKT 16
Db 117 hknkgkngkngkngwkt 132

```

```

RESULT 7
AAY93345
ID AAY93345 standard; peptide; 47 AA.
XX
AC AAY93345;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen fragment.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.

```

```

XX OS Synthetic.
XX OS Homo sapiens.
XX XX
PN WO200027415-A2.
XX XX
PD 18-MAY-2000.
XX XX
PF 09-NOV-1999; 99WO-US26377.
XX XX
PR 10-NOV-1998; 98US-0107844.
XX XX
PA (UTEM ) UNIV TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOSA/) MOUSA A S.
XX XX
PI Colman WR, Mousa AS;
XX XX
DR WPI; 2000-376306/32.
XX XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX XX
PS Claim 3; Page 36; 41pp; English.
XX XX
CC The present sequence represents a fragment of the light chain of human
CC high molecular weight kininogen. It is used to produce compounds of
CC the invention. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX XX
SQ Sequence 47 AA;

```

```

Query Match 94.9%; Score 93; DB 21; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HKNKGKNGKNGKNGKT 15
Db 33 hknkgkngkngkngwkt 47

```

```

RESULT 8
AAY93348
ID AAY93348 standard; peptide; 62 AA.
XX
AC AAY93348;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX

```


PF 09-NOV-1999; 99WO-US26377.
XX
XX 10-NOV-1998; 98US-0107844.
XX
XX (UTEM) UNIV TEMPLE.
PA (DUPONT) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
PI
XX
XX WPI; 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX
XX
XX Claim 6; Page 37; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
XX Sequence 62 AA;

Query Match 94.9%; Score 93; DB 21; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HKNKGKKNKGKNGWK 15
IIIIIIIIIIIIIIIIII
DB 48 hknkgkknkgkngwk 62

RESULT 9
AAR75186
ID AAR75186 standard; peptide; 63 AA.
XX
XX AAR75186;
XX
XX 05-DEC-1995 (first entry)
XX
XX Partial peptide of human HMW kininogen fragment 2.
XX
XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KW wound treating agent; bovine; growth promotion; fibroblast.
XX
XX Homo sapiens.
OS
XX JP07082172-A.
PN
XX 28-MAR-1995.
PD
XX 17-SEP-1993; 93JP-0230616.
PF
XX 17-SEP-1993; 93JP-0230616.
PR
XX 17-SEP-1993; 93JP-0230616.
XX
XX (FARH) HOECHST JAPAN KK.
PA
XX WPI; 1995-158909/21.
DR
XX A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX

PS Claim 8; Page 8; 8pp; Japanese.
XX
XX AAR75186 is a partial peptide corresponding to human kininogen
CC fragment 1, amino acids 458-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compens. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
XX
XX Sequence 63 AA;

Query Match 94.9%; Score 93; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HKNKGKKNKGKNGWK 15
IIIIIIIIIIIIIIIIII
DB 49 hknkgkknkgkngwk 63

RESULT 10
AAY93347
ID AAY93347 standard; peptide; 83 AA.
XX

AC AAY93347;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.
OS Homo sapiens.

PN WO200027415-A2.

PD 18-MAY-2000.

PF 09-NOV-1999; 99WO-US26377.

PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.
PA (DUPONT) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.

PI Colman WR, Mousa AS;

PI WPI; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration

PS Claim 5; Page 37; 41pp; English.

XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX

SQ Sequence 83 AA:

Query Match 94.9%; Score 93; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HKNKGKNGKHNGWK 15
|||||
Db 69 hknkgkngkhngwk 83

RESULT 11
AAR75181
ID AAR75181 standard; peptide; 131 AA.
XX
AC AAR75181;
XX
XX 05-DEC-1995 (first entry)
XX
XX Partial peptide of human HMW kininogen fragment 1.2.
DE
XX
KM high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KM wound treating agent; human; growth promotion; fibroblast.
XX
OS Homo sapiens.
XX
PN JP07082172-A.
XX
XX 28-MAR-1995.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX (FARM) HOECHST JAPAN KK.
XX
XX WPI; 1995-158909/21.
XX
XX A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
XX Claim 7; Page 7; 8pp; Japanese.
XX
XX AAR75181 is a partial peptide corresponding to human kininogen
CC fragment 1.2, amino acids 390-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 131 AA:

Query Match 94.9%; Score 93; DB 16; Length 131;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HKNKGKNGKHNGWK 15
|||||
Db 117 hknkgkngkhngwk 131

RESULT 12
AAR75195
ID AAR75195 standard; peptide; 12 AA.
XX
AC AAR75195;
XX
XX 16-OCT-2000 (first entry)
XX
XX Human high molecular weight kininogen domain 5 fragment #4.
DE
XX

KM Human; high molecular weight kininogen; HK;
KM two-chain high molecular weight kininogen; HKa;
KM angiogenesis inhibition; tumour; cancer; ocular disorder;
KM rheumatoid arthritis; endothelial cell apoptosis.
XX
OS Homo sapiens.
XX
XX WO200027866-A1.
XX
XX 18-MAY-2000.
XX
XX 05-NOV-1999; 99WO-US26419.
XX
XX 10-NOV-1998; 98US-0107833.
XX
XX (UTEM) UNIT TEMPLE.
XX
XX (MCCR/) MCCR/ R. K.
XX
XX MCCR/ R. K.
XX
XX WPI; 2000-376483/32.
XX
XX A pharmaceutical composition used to inhibit angiogenesis, inhibit
PT endothelial cell proliferation, and induce endothelial cell apoptosis
PT
XX
XX Claim 10; Page 29; 52pp; English.
XX
XX The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain
CC high molecular weight kininogen (HKa) by plasma kallikrein. HKa or a
CC synthetic compound comprising part or all of the present sequence may
CC be used in a pharmaceutical composition for inhibiting angiogenesis.
CC Angiogenesis occurs in a number of disease states, such as tumour
CC formation and expansion, and certain ocular disorders. It can also
CC occur in a rheumatoid joint, hastening joint destruction by allowing
CC an influx of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial
CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.
XX
SQ Sequence 12 AA:

Query Match 75.5%; Score 74; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 GKNGKHNGWK 16
|||||
Db 1 gkngkhngwk 12

RESULT 13
AAR75180
ID AAR75180 standard; peptide; 41 AA.
XX
AC AAR75180;
XX
XX 05-DEC-1995 (first entry)
XX
XX Partial peptide of HMW kininogen fragment 2.
XX
XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KM wound treating agent; bovine; growth promotion; fibroblast.
XX
XX Bos taurus.
XX
XX JP07082172-A.
XX
XX

PD 28-MAR-1995.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
PR 17-SEP-1993; 93JP-0230616.
XX
PA (FARRH) HOECHST JAPAN KK.
XX
DR WPI; 1995-158909/21.
XX
PT A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
PS Claim 6; Page 7; 8pp; Japanese.
XX
CC AAR75179 is a partial peptide corresponding to bovine kininogen
CC fragment 2, amino acids 456-496. Partial peptides of bovine and
CC human kininogen fragments 1,2, 1 and 2, are used in wound treating
CC agent compans. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 41 AA;
XX
Query Match 71.4%; Score 70; DB 16; Length 41;
Best Local Similarity 73.3%; Pred. No. 0.00076;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 HKNKGKNGKHNGWK 15
DB 27 hknkgknkgkhydw 41
XX
RESULT 14
AAR75178
ID AAR75178 standard; peptide; 110 AA.
XX
XX AAR75178;
AC
XX
DT 05-DEC-1995 (first entry)
XX
DE Partial peptide of HMW kininogen fragment 1.2.
XX
XX high molecular weight; kininogen; fragment; 1.2; 1: 2: partial;
KW wound treating agent; bovine; growth promotion; fibroblast.
XX
OS Bos taurus.
XX
FH Key Location/Qualifiers
FT Misc-difference 12 /label= Pro, Thr
FT Misc-difference 15 /label= Val or Ieu
FT Misc-difference 69 /label= Lys or Arg
FT
XX JPO7082172-A.
XX
XX 28-MAR-1995.
PD
XX
PF 17-SEP-1993; 93JP-0230616.
XX
PR 17-SEP-1993; 93JP-0230616.
XX
PA (FARRH) HOECHST JAPAN KK.
XX
DR WPI; 1995-158909/21.
XX
PT A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
PS Claim 4; Page 6; 8pp; Japanese.

XX
XX AAR75178 is a partial peptide corresponding to bovine kininogen
CC fragment 1.2, amino acids 387-496. Partial peptides of bovine and
CC human kininogen fragments 1,2, 1 and 2, are used in wound treating
CC agent compans. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 110 AA;
XX
Query Match 71.4%; Score 70; DB 16; Length 110;
Best Local Similarity 73.3%; Pred. No. 0.0021;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 HKNKGKNGKHNGWK 15
DB 96 hknkgknkgkhydw 110
XX
RESULT 15
AAR75178
ID AAR75178 standard; peptide; 20 AA.
XX
XX AAR75178;
AC
XX
DT 04-FEB-1997 (first entry)
XX
DE Human high polymer quininogen L-chain derived peptide.
XX
XX Human: high polymer; quininogen; L-chain; cell adhesion;
KW cancer metastasis; platelet aggregation; inhibition; wound;
KW inflammatory disease; arteriosclerosis; glomerular nephritis;
KW treatment.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Peptide 1..13 /note= "claimed peptide (claim 1)"
FT Peptide 13..20 /note= "claimed peptide (claim 6)"
FT
XX JPO8208692-A.
XX
XX 13-AUG-1996.
PD
XX
PF 28-SEP-1995; 95JP-0276418.
XX
PR 28-SEP-1994; 94JP-0259451.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 1996-421988/42.
XX
PT Cell adhesion inhibiting peptide(s), used as cancer metastasis
PT inhibitor - comprises partial amino acid sequence of human high
PT polymer quininogen L chain
XX
PS Claim 3; Page 2; 14pp; Japanese.
XX
CC The present peptide, and its claimed fragments, are derived from
CC residues 402-498 of the human high polymer quininogen L-chain. They
CC are useful in cell adhesion, cancer metastasis or platelet
CC aggregation inhibitors, and in wound, inflammatory disease,
CC arteriosclerosis or glomerular nephritis treating agents. The
CC present peptide was synthesised using a solid phase method, and
CC purified using a YMC-DOS-120A-S15/13 column.
XX
SQ Sequence 20 AA;
XX
Query Match 66.3%; Score 65; DB 17; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKNKGKNGKH 11
| | | | | | | | | |
Db 10 hknkgkngkh 20

Search completed: July 6, 2001, 09:09:19
Job time: 125 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:17 : Search time 113.68 Seconds
(without alignments)
14.932 Million cell updates/sec

Title: US-09-437-912-5

Perfect score: 167
Sequence: 1 HGHGQHGGLGHGKFKLDLLEHGGHV 28

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167	100.0	28	21	AA1981996
2	167	100.0	55	21	AA193346
3	167	100.0	62	21	AA193348
4	167	100.0	63	16	AA195186
5	167	100.0	83	21	AA193347
6	167	100.0	94	21	AA193351
7	167	100.0	131	16	AA195181
8	167	100.0	186	21	AA193345
9	167	100.0	255	21	AA193342
10	145	86.8	179	21	AA193353
11	98	58.7	47	21	AA193345

12	94	56.3	16	21	AA1981998
13	88	52.7	17	17	AA1907627
14	85	50.9	41	16	AA1975180
15	85	50.9	110	16	AA1975178
16	79	47.3	20	17	AA1907628
17	75	44.9	12	21	AA1981992
18	70	41.9	68	16	AA1975182
19	69	41.3	16	21	AA193350
20	69	41.3	186	21	AA193343
21	68	40.7	12	21	AA1981993
22	66	39.5	184	21	AA1982814
23	66	39.5	217	21	AA1982813
24	66	39.5	278	21	AA1982812
25	64	38.3	757	21	AA1931891
26	64	38.3	766	21	AA1931890
27	64	38.3	776	21	AA1931889
28	61	36.5	21	21	AA193344
29	60.5	36.2	330	21	AA1922665
30	60.5	36.2	330	21	AA1922665
31	60.5	36.2	344	21	AA1922664
32	60.5	36.2	344	21	AA1922664
33	60.5	36.2	398	21	AA1922663
34	60.5	36.2	398	21	AA1922663
35	60	35.9	20	17	AA1907626
36	57.5	34.4	28	21	AA1981997
37	57	34.1	337	19	AA1961155
38	56.5	33.8	69	16	AA1975179
39	55	32.9	792	22	AA1900027
40	54	32.3	309	21	AA1906065
41	54	32.3	309	21	AA1922955
42	54	32.3	389	21	AA1906064
43	54	32.3	389	21	AA1922954
44	54	32.3	425	21	AA1906063
45	54	32.3	453	21	AA1922953

ALIGNMENTS

RESULT 1
AA1981996
ID AAY81996 standard; peptide: 28 AA.
XX
AC AAY81996;
XX
DT 16-OCT-2000 (first entry)
XX
DE Human high molecular weight kininogen domain 5 fragment #5.
KW Human; high molecular weight kininogen; HK;
KW two-chain high molecular weight kininogen; HKa;
KW angiotensin inhibition; tumour; cancer; ocular disorder;
KW rheumatoid arthritis; endothelial cell apoptosis.
OS Homo sapiens.
PN WO200027866-A1.
PD 18-MAY-2000.
PF 05-NOV-1999; 99WO-US26419.
PR 10-NOV-1998; 98US-0107833.
XX
XX (UTEM) UNIV TEMPLE.
XX (MCCR/) MCCRAE R K.
XX
XX McCrae RK;
XX WPI: 2000-376483/32.
XX A pharmaceutical composition used to inhibit angiogenesis, inhibit
XX endothelial cell proliferation, and induce endothelial cell apoptosis
PT


```

PN WO200027415-A2.
PD 18-MAY-2000.
XX
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM ) UNIV TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
PI Colman WR, Mousa AS;
DR WPI: 2000-376306/32.
XX
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX
XX
PS Claim 8; Page 39; 41pp; English.
XX
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
XX
SQ Sequence 94 AA:

Query Match 100.0%; Score 167; DB 21; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.8e-17;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY 1 HGHEQOHGIGHGKFKPLDDLEHOGGHV 28
| | | | | | | | | | | | | | | | | | | | |
Db 26 hgneqnglignhktloddlehgghv 53

RESULT 7
AAR75181
ID AAR75181 standard; peptide: 131 AA.
XX
XX AAR75181;
XX
XX 05-DEC-1995 (first entry)
DE Partial peptide of human HMW kininogen fragment 1.2.
XX
XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KW wound treating agent; human; growth promotion; fibroblast.
XX
XX Homo sapiens.
XX
XX JP07082172-A.
XX
XX 28-MAR-1995.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX (FARH ) HOECHST JAPAN KK.
XX
XX WPI: 1995-156909/21.

```

XX A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
XX
PS Claim 7; Page 7; 8pp; Japanese.
XX
CC AAR5181 is a partial peptide corresponding to human kininogen
CC fragment 1.2, amino acids 390-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compsns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 131 AA:

Query Match 100.0%; Score 167; DB 16; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGEEOHGLGHGHRKRLDDLEHOGSHV 28
|||||
Db 74 hghgqghlyghghkrloddlehggshv 101

RESULT 8
AA93349
ID AAY93349 standard; peptide: 186 AA:
XX
XX AAY93349;
AC
XX
DT 04-SEP-2000 (first entry)
DE
XX Light chain of human high molecular weight kininogen analogue.
XX
XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
KW
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
XX
PF 09-NOV-1999; 99WO-US26377.
XX
XX 10-NOV-1998; 98US-0107844.
PR
XX (UTEM) UNIT TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COIM/) COLMAN W R.
PA (MOSA/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
PI
XX WPI; 2000-376306/32.
DR
XX
XX Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
XX
XX Claim 9; Page 38; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for

CC Inhibiting endothelial cell proliferation. The compounds can also be
 CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.
 XX

SQ Sequence 186 AA;

Query Match 100.0%; Score 167; DB 21; Length 186;
 Best Local Similarity 100.0%; Pred. No. 1,7e-16;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEOQHGLGHGKFKLDDLEHGGHV 28
 |||||
 Db 5 hgheqghlgghkfklddlehgghv 32

RESULT 9

AAV93342
 ID AAV93342 standard; protein; 255 AA.

AC AAV93342;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Homo sapiens.

PN WO200027415-A2.

PD 18-MAY-2000.

PF 09-NOV-1999; 99WO-US26377.

PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

PI Colman WR, Mousa AS;

DR WPI; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
 PT that inhibit endothelial cell migration -

PS Disclosure: Page 3; 41pp; English.

XX The present sequence represents the light chain of human high molecular
 CC weight kininogen. High molecular weight kininogen is a 120 kDa
 CC glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method
 CC of the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be
 CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.

SQ Sequence 255 AA;

Query Match 100.0%; Score 167; DB 21; Length 255;
 Best Local Similarity 100.0%; Pred. No. 2,4e-16;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEOQHGLGHGKFKLDDLEHGGHV 28
 |||||
 Db 74 hgheqghlgghkfklddlehgghv 101

RESULT 10

AAV93353
 ID AAV93353 standard; peptide; 179 AA.

AC AAV93353;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.

PN WO200027415-A2.

PD 18-MAY-2000.

PF 09-NOV-1999; 99WO-US26377.

PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

PI Colman WR, Mousa AS;

DR WPI; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
 PT that inhibit endothelial cell migration -

PS Claim 11; Page 40-41; 41pp; English.

XX The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a
 CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method
 CC of the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be
 CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.

SQ Sequence 179 AA;

Query Match 86.8%; Score 145; DB 21; Length 179;
 Best Local Similarity 100.0%; Pred. No. 2,3e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 EGOHGLGHGKFKLDDLEHGGHV 28
 |||||
 Db 1 eqghlgghkfklddlehgghv 25

RESULT 11

AAV93345
 ID AAV93345 standard; peptide; 47 AA.

```
XX AAY93345;
AC
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen fragment.
XX
KM Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KM plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KM endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
PN WO200027415-A2.
PD 18-MAY-2000.
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM ) UNIV TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
DR WPI; 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
PS Claim 3; Page 36; 41pp; English.
XX
CC The present sequence represents a fragment of the light chain of human
CC high molecular weight kininogen. It is used to produce compounds of
CC the invention. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells.
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 47 AA;
OY
Y 12 GHKFKLDDLEHGGHV 28
Db 1 ghkfklddlehgghv 17
RESULT 12
AAV81998
ID AAV81998 standard; peptide; 16 AA.
XX
AC AAV81998;
XX
DT 16-OCT-2000 (first entry)
XX
DE Human two-chain high molecular weight kininogen domain 5 fragment #7.
XX
KW Human; high molecular weight kininogen; HK;
```

```
KM two-chain high molecular weight kininogen; HKa;
KM angiogenesis inhibition; tumour; cancer; ocular disorder;
KM rheumatoid arthritis; endothelial cell apoptosis.
XX
XX Homo sapiens.
XX
PN WO200027866-A1.
PD 18-MAY-2000.
PF 05-NOV-1999; 99WO-US26419.
XX
PR 10-NOV-1998; 98US-0107833.
XX
PA (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI McCrae RK;
DR WPI; 2000-376483/32.
XX
PT A pharmaceutical composition used to inhibit angiogenesis, inhibit
PT endothelial cell proliferation, and induce endothelial cell apoptosis
XX
PS Claim 9; Page 28; 52pp; English.
XX
CC The present sequence is derived from human two-chain high molecular
CC weight kininogen (Hka) domain 5. Hka is product of high molecular
CC weight kininogen (HK) cleavage by plasma kallikrein. HK is a 120 kD
CC glycoprotein which binds with high affinity to endothelial cells. Hka
CC or a synthetic compound comprising the present sequence may be
CC used in a pharmaceutical composition for inhibiting angiogenesis.
CC Angiogenesis occurs in a number of disease states, such as tumour
CC formation and expansion, and certain ocular disorders. It can also occur
CC in a rheumatoid joint, hastening joint destruction by allowing an influx
CC of leukocytes. The composition may inhibit angiogenesis by inhibiting
CC endothelial cell proliferation or by inducing endothelial cell
CC apoptosis. Peptides used in the composition may be recombinant peptides,
CC natural peptides, or synthetic peptides. They may also be chemically
CC synthesised, using, for example, solid phase synthesis methods.
XX
SQ Sequence 16 AA;
OY
Y 12 GHKFKLDDLEHGGH 27
Db 1 ghkfklddlehggh 16
RESULT 13
AAW07627
ID AAW07627 standard; peptide; 17 AA.
XX
AC AAW07627;
XX
DT 04-FEB-1997 (first entry)
XX
DE Human high polymer quininogen L-chain derived peptide.
XX
KW Human; high polymer; quininogen; L-chain.
XX
OS Homo sapiens.
XX
PN JP08208692-A.
PD 13-AUG-1996.
XX
PF 28-SEP-1995; 95JP-0276418.
```

XX 28-SEP-1994; 94JP-0259451.
 XX (SUMU) SUMITOMO SEIYAKU KK.
 PA
 XX
 DR WPI; 1996-421988/42.
 XX
 PT Cell adhesion inhibiting peptide(s), used as cancer metastasis
 PT inhibitor - comprises partial amino acid sequence of human high
 PT polymer quininogen L chain
 XX
 PS Example; Page 8; 14pp; Japanese.
 XX
 CC The present peptide is derived from residues 402-498 of the human
 CC high polymer quininogen L-chain. It was synthesised using a solid
 CC phase method, and purified using a YMC-DOS-120A-S15/13 column.
 XX
 SQ Sequence 17 AA;

Query Match 52.7%; Score 88; DB 17; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEOQHGLGHGHR 14
 |||||
 DB 4 hheqghqghghghk 17

RESULT 14

AAR75180
 ID AAR75180 standard; peptide; 41 AA.

AC AAR75180;

DT 05-DEC-1995 (first entry)

DE Partial peptide of HMW kininogen fragment 2.

KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 KM wound treating agent; bovine; growth promotion; fibroblast.

OS Bos taurus.

PN JP07082172-A.

PD 28-MAR-1995.

PF 17-SEP-1993; 93JP-0230616.

PR 17-SEP-1993; 93JP-0230616.

PA (FARH) HOECHST JAPAN KK.

DR WPI; 1995-158909/21.

PT A wound treating agent contg. a partial peptide of kininogen -
 PT have growth promotion activity of fibroblasts.

PS Claim 6; Page 7; 8pp; Japanese.

CC AAR75179 is a partial peptide corresponding to bovine kininogen
 CC fragment 2, amino acids 456-496. Partial peptides of bovine and
 CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
 CC agent compns. and act as the active component. The fragments are
 CC useful in wound treating because they have growth promotion activity
 CC on fibroblasts.

SQ Sequence 41 AA;

Query Match 50.9%; Score 85; DB 16; Length 41;
 Best Local Similarity 55.6%; Pred. No. 1.7e-05;

Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;
 OY 1 HGHEOQHGLGHGHRFLDDLEHOGH 27
 |||::|||
 DB 6 hghqkqhghghghk-----hghgh 24

RESULT 15

AAR75178
 ID AAR75178 standard; peptide; 110 AA.

AC AAR75178;

DT 05-DEC-1995 (first entry)

DE Partial peptide of HMW kininogen fragment 1.2.

KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
 KM wound treating agent; bovine; growth promotion; fibroblast.

OS Bos taurus.

FT Key Location/Qualifiers

FT Misc-difference 12 /label= Pro, Thr

FT Misc-difference 15 /label= Val or Leu

FT Misc-difference 69 /label= Lys or Arg

PN JP07082172-A.

PD 28-MAR-1995.

PF 17-SEP-1993; 93JP-0230616.

PR 17-SEP-1993; 93JP-0230616.

PA (FARH) HOECHST JAPAN KK.

DR WPI; 1995-158909/21.

PT A wound treating agent contg. a partial peptide of kininogen -
 PT have growth promotion activity of fibroblasts.

PS Claim 4; Page 6; 8pp; Japanese.

CC AAR75178 is a partial peptide corresponding to bovine kininogen
 CC fragment 1.2, amino acids 387-496. Partial peptides of bovine and
 CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
 CC agent compns. and act as the active component. The fragments are
 CC useful in wound treating because they have growth promotion activity
 CC on fibroblasts.

SQ Sequence 110 AA;

Query Match 50.9%; Score 85; DB 16; Length 110;
 Best Local Similarity 55.6%; Pred. No. 5.3e-05;
 Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

OY 1 HGHEOQHGLGHGHRFLDDLEHOGH 27
 |||::|||
 DB 75 hghqkqhghghghk-----hghgh 93

Search completed: July 6, 2001, 09:09:18
 Job time: 124 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:18 ; Search time 113.68 seconds

(without alignments)
8.533 Million cell updates/sec

Title: US-09-437-912-7

Perfect score: 94

Sequence: 1 GHKFKLDLDLEHGGH 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing First 45 summaries

Database : A.Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	16	AA1981998	Human two-chain hi
2	94	100.0	28	AA1981996	Human high molecu
3	94	100.0	47	AA1981996	Human high molecu
4	94	100.0	55	AA1981996	Human high molecu
5	94	100.0	62	AA1981996	Human high molecu
6	94	100.0	63	AA1981996	Human high molecu
7	94	100.0	83	AA1981996	Human high molecu
8	94	100.0	94	AA1981996	Human high molecu
9	94	100.0	131	AA1981996	Human high molecu
10	94	100.0	179	AA1981996	Human high molecu
11	94	100.0	186	AA1981996	Human high molecu

12	94	100.0	255	21	AA1981998	Human high molecu
13	75	79.8	20	17	AA1981998	Human high molecu
14	64	68.1	12	21	AA1981998	Human high molecu
15	45	47.9	631	21	AA1981998	Human high molecu
16	43	45.7	57	21	AA1981998	Human high molecu
17	43	45.7	107	18	AA1981998	Human high molecu
18	43	45.7	107	18	AA1981998	Human high molecu
19	43	45.7	107	18	AA1981998	Human high molecu
20	43	45.7	107	20	AA1981998	Human high molecu
21	43	45.7	107	20	AA1981998	Human high molecu
22	43	45.7	135	18	AA1981998	Human high molecu
23	43	45.7	135	20	AA1981998	Human high molecu
24	43	45.7	135	22	AA1981998	Human high molecu
25	41	43.6	477	21	AA1981998	Human high molecu
26	41	43.6	951	21	AA1981998	Human high molecu
27	41	43.6	1477	16	AA1981998	Human high molecu
28	41	43.6	1477	18	AA1981998	Human high molecu
29	41	43.6	1477	20	AA1981998	Human high molecu
30	40	42.6	46	21	AA1981998	Human high molecu
31	40	42.6	184	21	AA1981998	Human high molecu
32	40	42.6	195	20	AA1981998	Human high molecu
33	40	42.6	217	21	AA1981998	Human high molecu
34	40	42.6	238	22	AA1981998	Human high molecu
35	40	42.6	248	21	AA1981998	Human high molecu
36	40	42.6	278	21	AA1981998	Human high molecu
37	40	42.6	317	19	AA1981998	Human high molecu
38	40	42.6	317	19	AA1981998	Human high molecu
39	40	42.6	413	21	AA1981998	Human high molecu
40	40	42.6	606	20	AA1981998	Human high molecu
41	40	42.6	1004	20	AA1981998	Human high molecu
42	40	42.6	1004	20	AA1981998	Human high molecu
43	40	42.6	1025	16	AA1981998	Human high molecu
44	40	42.6	1035	20	AA1981998	Human high molecu
45	40	42.6	1079	20	AA1981998	Human high molecu

ALIGNMENTS

RESULT 1	AA1981998	standard; peptide; 16 AA.	Light chain of hum
ID	AA1981998	standard; peptide; 16 AA.	Human high molecu
AC	AA1981998	standard; peptide; 16 AA.	Human high molecu
DT	AA1981998	standard; peptide; 16 AA.	Human high molecu
DE	AA1981998	standard; peptide; 16 AA.	Human high molecu
KW	AA1981998	standard; peptide; 16 AA.	Human high molecu
OS	AA1981998	standard; peptide; 16 AA.	Human high molecu
PN	AA1981998	standard; peptide; 16 AA.	Human high molecu
PD	AA1981998	standard; peptide; 16 AA.	Human high molecu
PF	AA1981998	standard; peptide; 16 AA.	Human high molecu
PK	AA1981998	standard; peptide; 16 AA.	Human high molecu
PA	AA1981998	standard; peptide; 16 AA.	Human high molecu
PI	AA1981998	standard; peptide; 16 AA.	Human high molecu
PT	AA1981998	standard; peptide; 16 AA.	Human high molecu

CC Inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.

XX SQ Sequence 47 AA;

Query Match 100.0%; Score 94; DB 21; Length 47;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKFKLDDLEHOGH 16
|
DB 1 ghkfklddlehggh 16

RESULT 4

AAV93346
ID AAV93346 standard; peptide; 55 AA.

XX AC AAV93346;

XX DT 04-SEP-2000 (first entry)

XX DE Light chain of human high molecular weight kininogen analogue.

XX KM Human; high molecular weight kininogen; glycoprotein; endothelial cell;
XX KW Plasma kallikrein; heavy chain; analogue; angiogenesis;
XX KM endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.

OS Homo sapiens.

XX PN WO200027415-A2.

XX PD 18-MAY-2000.

XX PF 09-NOV-1999; 99WO-US26377.

XX PR 10-NOV-1998; 98US-0107844.

XX PA (UTEM) UNIV TEMPLE.

XX PA (DUPO) DUPONT PHARM CO.

XX PA (COLM/) COLMAN W R.

XX PA (MOUS/) MOUSA A S.

XX PI Colman WR, Mousa AS;

XX DR WPI; 2000-376306/32.

XX PT Method for inhibiting endothelial cell proliferation, using compound
XX PT that inhibit endothelial cell migration

XX PS Claim 4; Page 36; 41pp; English.

XX CC The present sequence represents an analogue of the light chain of human
XX CC high molecular weight kininogen. High molecular weight kininogen is a
XX CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
XX CC where it is cleaved by plasma kallikrein into heavy and light chains.
XX CC Analogues of high molecular weight kininogen are used in the method
XX CC of the invention. The specification describes a method of inhibiting
XX CC endothelial cell proliferation. The method comprises contacting
XX CC endothelial cells with a compound containing high molecular weight
XX CC kininogen analogues. The method and the compounds can be used for
XX CC inhibiting endothelial cell proliferation. The compounds can also be
XX CC used for inhibiting angiogenesis. The compounds can also be used to
XX CC inhibit migration of endothelial cells to vitronectin.

XX SQ Sequence 55 AA;

Query Match 100.0%; Score 94; DB 21; Length 55;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKFKLDDLEHOGH 16
|
DB 37 ghkfklddlehggh 52

RESULT 5

AAV93348
ID AAV93348 standard; peptide; 62 AA.

XX AC AAV93348;

XX DT 04-SEP-2000 (first entry)

XX DE Light chain of human high molecular weight kininogen analogue.

XX KM Human; high molecular weight kininogen; glycoprotein; endothelial cell;
XX KW Plasma kallikrein; heavy chain; analogue; angiogenesis;
XX KM endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.

OS Homo sapiens.

XX PN WO200027415-A2.

XX PD 18-MAY-2000.

XX PF 09-NOV-1999; 99WO-US26377.

XX PR 10-NOV-1998; 98US-0107844.

XX PA (UTEM) UNIV TEMPLE.

XX PA (DUPO) DUPONT PHARM CO.

XX PA (COLM/) COLMAN W R.

XX PA (MOUS/) MOUSA A S.

XX PI Colman WR, Mousa AS;

XX DR WPI; 2000-376306/32.

XX PT Method for inhibiting endothelial cell proliferation, using compound
XX PT that inhibit endothelial cell migration

XX PS Claim 6; Page 37; 41pp; English.

XX CC The present sequence represents an analogue of the light chain of human
XX CC high molecular weight kininogen. High molecular weight kininogen is a
XX CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
XX CC where it is cleaved by plasma kallikrein into heavy and light chains.
XX CC Analogues of high molecular weight kininogen are used in the method
XX CC of the invention. The specification describes a method of inhibiting
XX CC endothelial cell proliferation. The method comprises contacting
XX CC endothelial cells with a compound containing high molecular weight
XX CC kininogen analogues. The method and the compounds can be used for
XX CC inhibiting endothelial cell proliferation. The compounds can also be
XX CC used for inhibiting angiogenesis. The compounds can also be used to
XX CC inhibit migration of endothelial cells to vitronectin.

XX SQ Sequence 62 AA;

Query Match 100.0%; Score 94; DB 21; Length 62;
Best Local Similarity 100.0%; Pred. No. 5.6e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKFKLDDLEHOGH 16
|
DB 16 ghkfklddlehggh 31

RESULT 6
AAR75186

XX Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
PS Claim 8; Page 39; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 94 AA;

Query Match 100.0%; Score 94; DB 21; Length 94;
Best Local Similarity 100.0%; Pred. No. 9e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKFKLDDLEHOGGH 16
DB 37 gnhfklddlehgh 52

RESULT 9
AAR75181
ID AAR75181 standard; peptide; 131-AA.
XX
AC AAR75181;
XX
DT 05-DEC-1995 (first entry)
XX
DE Partial peptide of human HMW kininogen fragment 1.2.
XX
KW high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KW wound treating agent; human; growth promotion; fibroblast.
XX
OS Homo sapiens.
XX
PN JP07082172-A.
XX
PD 28-MAR-1995;
XX
PF 17-SEP-1993; 93JP-0230616.
XX
PR 17-SEP-1993; 93JP-0230616.
XX
PA (FARNH) HOECHST JAPAN KK.
XX
DR WPI; 1995-158909/21.
XX
PT A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
PS Claim 7; Page 7; 8pp; Japanese.
XX
CC AAR75181 is a partial peptide corresponding to human kininogen
CC fragment 1.2, amino acids 390-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 131 AA;

Query Match 100.0%; Score 94; DB 16; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKFKLDDLEHOGGH 16
DB 85 gnhfklddlehgh 100

RESULT 10
AAR93353
ID AAR93353 standard; peptide; 179 AA.
XX
AC AAR93353;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
XX
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
XX
DR WPI; 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
PS Claim 11; Page 40-41; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 179 AA;

Query Match 100.0%; Score 94; DB 21; Length 179;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GHKFKLDDLEHOGGH 16
DB 9 gnhfklddlehgh 24

RESULT 11
AA93349
ID AAY93349 standard; peptide; 186 AA.
XX
AC AAY93349;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PT (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
PI
DR WPI; 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
PS Claim 9; Page 38; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 186 AA;

Query Match 100.0%; Score 94; DB 21; Length 186;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHFKLDDDLHOGH 16
DB 16 ghfkldddlehggh 31

RESULT 12
AA93342
ID AAY93342 standard; protein; 255 AA.
XX
AC AAY93342;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen.

XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Homo sapiens.
XX
PN WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PT (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
PI
DR WPI; 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration -
XX
PS Disclosure; Page 3; 41pp; English.
XX
CC The present sequence represents the light chain of human high molecular
CC weight kininogen. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 255 AA;

Query Match 100.0%; Score 94; DB 21; Length 255;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHFKLDDDLHOGH 16
DB 85 ghfkldddlehggh 100

RESULT 13
AAW07628
ID AAW07628 standard; peptide; 20 AA.
XX
AC AAW07628;
XX
DT 04-FEB-1997 (first entry)
XX
DE Human high polymer quininogen L-chain derived peptide.
XX
KW Human; high polymer; quininogen; L-chain.
XX
OS Homo sapiens.
XX
PN JP08208692-A.
XX
PD 13-AUG-1996.
XX
PF 28-SEP-1995; 95JP-0276418.

```

XX 28-SEP-1994; 94JP-0259451.
XX
XX (SUMO ) SUMITOMO SEIYAKU KK.
XX
XX WPI; 1996-421988/42.
XX
XX Cell adhesion inhibiting peptide(s), used as cancer metastasis
XX inhibitor - comprises partial amino acid sequence of human high
XX polymer quininogen L chain
XX
XX Example; Page 8; 14pp; Japanese.
XX
XX The present peptide is derived from residues 402-498 of the human
XX high polymer quininogen L-chain. It was synthesised using a solid
XX phase method, and purified using a YMC-DOS-120A-S15/13 column.
XX
XX Sequence 20 AA;
SQ

Query Match          79.8%; Score 75; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FKLDLLEHOGGH 16
   |||||
   1 FKlddlehggh 13

Db

RESULT 14
AA79193
ID AAY81993 standard; peptide; 12 AA.
XX
XX AAY81993;
XX
XX 16-OCT-2000 (first entry)
XX
XX Human high molecular weight kininogen domain 5 fragment #2.
XX
XX Human; high molecular weight kininogen; HK;
XX two-chain high molecular weight kininogen; HKA;
XX angiotensin inhibition; tumour; cancer; ocular disorder;
XX rheumatoid arthritis; endothelial cell apoptosis.
XX
XX Homo sapiens.
XX
XX WO200027866-A1.
XX
XX 18-MAY-2000.
XX
XX 05-NOV-1999; 99WO-US26419.
XX
XX 10-NOV-1998; 98US-0107833.
XX
XX (UTEM ) UNIV TEMPLE.
XX (MCCR/) MCCRAE R K.
XX
XX MCCRAE RK;
XX
XX WPI; 2000-376483/32.
XX
XX A pharmaceutical composition used to inhibit angiogenesis, inhibit
XX endothelial cell proliferation, and induce endothelial cell apoptosis
XX
XX Claim 5; Page 28; 52pp; English.
XX
XX The present sequence is derived from human high molecular
XX weight kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds
XX with high affinity to endothelial cells, where it is cleaved to
XX two-chain high molecular weight kininogen (HKA) by plasma kallikrein.
XX HKA or a synthetic compound comprising part or all of the present
XX sequence may be used in a pharmaceutical composition for inhibiting

```

```

CC angiogenesis. Angiogenesis occurs in a number of disease states, such
CC as tumour formation and expansion, and certain ocular disorders. It can
CC also occur in a rheumatoid joint, hastening joint destruction by
CC allowing an influx of leukocytes. The composition may inhibit
CC angiogenesis by inhibiting endothelial cell proliferation or by
CC inducing endothelial cell apoptosis. Peptides used in the composition
CC may be recombinant peptides, natural peptides, or synthetic peptides.
CC They may also be chemically synthesised, using, for example, solid
CC phase synthesis methods.
XX
XX Sequence 12 AA;
SQ

Query Match          68.1%; Score 64; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LDLDLLEHOGGH 16
   |||||
   1 ldldlehggh 11

Db

RESULT 15
AA79183
ID AAY79183 standard; Protein; 631 AA.
XX
XX AAY79183;
XX
XX 19-JUN-2000 (first entry)
XX
XX Haematopoietic stem cell specific protein.
XX
XX Haematopoietic stem cell; immune system disorder;
XX leukaemia; anti-leukaemic; immunomodulator; therapy; mouse.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX MISC-difference 497 /note= "encoded by TTA"
XX FT MISC-difference 523 /note= "encoded by CAT"
XX FT MISC-difference 548 /note= "encoded by TAA"
XX FT MISC-difference 548 /note= "encoded by TAA"
XX
XX WO200011168-A2.
XX
XX 02-MAR-2000.
XX
XX 20-AUG-1999; 99WO-US19052.
XX
XX 21-AUG-1998; 98US-0138132.
XX
XX (UYPR-) UNIV PRINCETON.
XX
XX Lemischka I, Moore K;
XX
XX WPI; 2000-237650/20.
XX
XX N-PSDB; AA294124.
XX
XX Hematopoietic stem cell signaling proteins modulating replication and
XX differentiation for treating immune system disorders and leukaemia -
XX
XX Claim 21; Page 229-231; 256pp; English.
XX
XX The present sequence is that of a mouse haematopoietic stem cell
XX (HSC) specific protein. It is an example of claimed HSC-specific
XX proteins (see AAY79176-93) predicted from novel isolated HSC-specific
XX nucleic acids (see AA294077-131). The HSCs are especially primitive
XX HSCs (PHSCs) such as umbilical cord cells, bone marrow cells and
XX foetal liver cells. The encoded proteins are growth factors,
XX transcription factors, splicing factors, capping factors, transport
XX proteins, translation factors or replication factors that modulate

```

CC HSC activity, especially differentiation or replication. The
 CC invention provides claimed methods: for identifying PHSC-specific
 CC nucleic acids; for generating a stem cell/progenitor cell from
 CC PHSCs; for identifying the presence of a PHSC in a sample; for
 CC identifying the presence in a sample of a compound that modulates
 CC HSC activity; for using such a compound to treat an immune system
 CC condition, especially leukaemia; for introducing exogenous nucleic
 CC acid into a HSC; and for ex vivo expansion of HSCs. Also claimed
 CC are vectors, host cells, and an antibody that specifically binds a
 CC an HSC-specific protein.

XX
 SQ Sequence 631 AA;

Query Match 47.9%; Score 45; DB 21; Length 631;
 Best local Similarity 50.0%; Pred. NO. 18;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 GHKFKLDDLEHOGH 16
 ||:|| | : ||
 Db 351 ghfrkrsdickkqgh 366

Search completed: July 6, 2001, 09:09:18
 Job time: 124 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:23 ; Search time 56.74 Seconds
(without alignments)
5.681 Million cell updates/sec

Title: US-09-437-912-7

Perfect score: 94
Sequence: 1 GHKFKLDLDEHOGCH 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 2014635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	45.7	106	2	US-08-820-825-14
2	43	45.7	106	4	US-09-307-817-14
3	43	45.7	107	1	US-08-409-731A-2
4	43	45.7	107	2	US-08-470-298B-2
5	43	45.7	107	2	US-09-023-073A-2
6	43	45.7	135	2	US-08-820-825-2
7	43	45.7	135	3	US-08-899-031-1
8	43	45.7	135	3	US-09-307-817-2
9	41	43.6	1477	3	US-08-492-459-10
10	41	43.6	1477	3	US-08-423-752-10
11	41	43.6	1477	3	US-08-945-994-3
12	41	43.6	1477	4	US-08-716-873-24
13	40	42.6	1079	3	US-09-136-652-2
14	39	41.5	1001	4	US-09-060-410-2
15	38	40.4	463	1	US-08-679-635A-4
16	36.5	38.8	509	1	US-08-030-096-2
17	36	38.3	344	1	US-08-403-866-9
18	36	38.3	393	4	US-09-127-124-1
19	36	38.3	481	5	PCR-US93-07213-9
20	36	38.3	604	2	US-08-468-576B-12
21	36	38.3	604	2	US-08-468-579B-12
22	36	38.3	604	3	US-08-468-577B-12
23	36	38.3	714	5	PCR-US93-07213-5
24	36	38.3	758	2	US-08-222-617A-6
25	36	38.3	858	5	PCR-US93-07213-2
26	36	38.3	3665	2	US-08-222-617A-13
27	36	38.3	3666	2	US-08-222-617A-12

28	36	38.3	3712	2	US-08-222-617A-4	Sequence 4, Appl
29	36	38.3	3712	2	US-08-222-617A-25	Sequence 25, Appl
30	36	38.3	3727	2	US-08-222-617A-27	Sequence 27, Appl
31	36	38.3	3778	2	US-08-222-617A-2	Sequence 2, Appl
32	35	37.2	86	4	US-08-905-223-420	Sequence 420, App
33	35	37.2	123	2	US-08-937-972-1	Sequence 1, Appl
34	35	37.2	447	1	US-08-476-008-67	Sequence 67, Appl
35	35	37.2	447	1	US-08-306-063-67	Sequence 67, Appl
36	35	37.2	447	1	US-08-833-485-67	Sequence 67, Appl
37	35	37.2	532	2	US-08-899-324-33	Sequence 33, Appl
38	35	37.2	532	2	US-08-329-892B-33	Sequence 33, Appl
39	35	37.2	583	2	US-08-616-392C-4	Sequence 4, Appl
40	35	37.2	788	2	US-08-907-166-6	Sequence 6, Appl
41	35	37.2	831	2	US-09-047-026A-4	Sequence 4, Appl
42	34	36.2	50	4	US-08-844-188-5	Sequence 5, Appl
43	34	36.2	123	4	US-08-844-188-36	Sequence 36, Appl
44	34	36.2	123	4	US-08-844-188-41	Sequence 41, Appl
45	34	36.2	154	1	US-08-102-942A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-820-825-14
; Sequence 14, Application US/08820825
; Patent No. 5945309
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENTZ, REINER L.
; APPLICANT: DILLON, PATRICK
; TITLE OF INVENTION: CYTOSTATIN III
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,825
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36, 373
; REFERENCE/DOCKET NUMBER: PF222
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-820-825-14

Query Match 45.7%; Score 43; DB 2; Length 106;
Best Local Similarity 58.3%; Pred. No. 2.8;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 5 KLDDLEHOGCH 16
| |::| | | |

Db 37 KPDKIEHOGNH 48

RESULT 2
US-09-307-817-14
Sequence 14, Application US/09307817
Patent No. 6232291
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER L.
TITLE OF INVENTION: CYTOSTATIN III
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/307,817
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/820,825
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF222
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-307-817-14

Query Match 45.7%; Score 43; DB 4; Length 106;
Best Local Similarity 58.3%; Pred. No. 2.8;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDLEHOGGH 16
| | : : | | | | |
| | : : | | | | |
DB 37 KPDKIEHOGNH 48

RESULT 3
US-08-409-731A-2
Sequence 2, Application US/08409731A
Patent No. 5658758
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: CYTOSTATIN I
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/409,731A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Robert H
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PF175
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8512
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-409-731A-2

Query Match 45.7%; Score 43; DB 1; Length 107;
Best Local Similarity 58.3%; Pred. No. 2.9;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 KLDDLEHOGGH 16
| | : : | | | | |
| | : : | | | | |
DB 38 KPDKIEHOGNH 49

RESULT 4
US-08-470-298B-2
Sequence 2, Application US/08470298B
Patent No. 5844081
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: GENTZ, REINER
APPLICANT: YU, GUO-LIANG
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: CYTOSTATIN I
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,298B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ALLAN A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF175D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8512
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-470-298B-2

Query Match 45.7%; Score 43; DB 2; Length 107;
Best Local Similarity 58.3%; Pred. No. 2.9;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KLDDLEHOGGH 16
| |:::||||| |
Db 38 KPDKETEHOGNH 49

RESULT 5
US-09-023-073A-2
Sequence 2, Application US/09023073A
Patent No. 5977309
GENERAL INFORMATION:
APPLICANT: NI, Jian
APPLICANT: Gentz, Reiner
APPLICANT: Yu, Guo-Liang
APPLICANT: Rosen, Craig A
TITLE OF INVENTION: Cytostatin I
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MARYLAND
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,073A
FILING DATE: 13-FEB-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maies, Michele M.
REGISTRATION NUMBER: P-43,975
REFERENCE/DOCKET NUMBER: PF175D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5772
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-023-073A-2

Query Match 45.7%; Score 43; DB 2; Length 107;
Best Local Similarity 58.3%; Pred. No. 2.9;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KLDDLEHOGGH 16
| |:::||||| |
Db 38 KPDKETEHOGNH 49

RESULT 6
US-08-820-825-2
Sequence 2, Application US/08820825
Patent No. 5945309

GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER L.
APPLICANT: DILLON, PATRICK
TITLE OF INVENTION: CYTOSTATIN III
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/820,825
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF222
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-820-825-2

Query Match 45.7%; Score 43; DB 2; Length 135;
Best Local Similarity 58.3%; Pred. No. 3.7;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KLDDLEHOGGH 16
| |:::||||| |
Db 38 KPDKETEHOGNH 49

RESULT 7
US-08-899-031-1
Sequence 1, Application US/08899031
Patent No. 6046027
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Guegler, Karl J.
APPLICANT: Shah, Purvu
TITLE OF INVENTION: HUMAN RETINOID BINDING PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTESQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,031
FILING DATE: Herewith

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0349 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAT01
CLONE: 367304
US-08-899-031-1

Query Match 45.7%; Score 43; DB 3; Length 135;
Best Local Similarity 58.3%; Pred. No. 3.7;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KLDDLEHOGGH 16
| | : | | | | |
Db 38 KPKEIEHOGNH 49

RESULT 8
US-09-307-817-2
Sequence 2, Application US/09307817
Patent No. 6232291
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, RETNER L.
APPLICANT: DILLON, PATRICK
TITLE OF INVENTION: CYTOSTATIN III
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/307,817
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/820,825
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF222
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-307-817-2

Query Match 45.7%; Score 43; DB 4; Length 135;
Best Local Similarity 58.3%; Pred. No. 3.7;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KLDDLEHOGGH 16
| | : | | | | |
Db 38 KPKEIEHOGNH 49

RESULT 9
US-08-492-459-10
Sequence 10, Application US/08492459
Patent No. 6015689
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: REGULATION OF AUREOBASIDIUM SENSITIVITY IN FUNGUS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette, 3.5 inch, 1.4 mb
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/492,459
FILING DATE: June 20, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/243,403
FILING DATE: May 16, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1477
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-492-459-10

Query Match 43.6%; Score 41; DB 3; Length 1477;
Best Local Similarity 46.7%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 HKFKLDDLEHOGGH 16
| | | | : | | | |
Db 1352 HKFKLDDLEHOGGH 1366

RESULT 10
US-08-423-752-10
Sequence 10, Application US/08423752
Patent No. 6022949

Fri Jul 6 09:48:32 2001

us-09-437-912-7.rai

Page 7

	Matches	9;	Conservative	0;	Mismatches	2;	Indels	2;	Gaps	1;
QY	1	GHKKKIDDDLEHQ	13							
Db	118	GH--KLDDTDEDQ	128							

Search completed: July 6, 2001, 09:10:23
Job time: 189 sec

1

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:18:01 ; Search time 73.59 Seconds

(without alignments)
16.562 Million cell updates/sec

Title: US-09-437-912-7

Sequence: 1 GHKFKLDLDLEHGGH 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR_66:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	644	1 KGHUHI	kininogen, HMW pre
2	47	50.0	414	2 H64231	protein L homolog
3	44	46.8	541	2 T34850	probable acid--CoA
4	43	45.7	296	2 A41730	nucleophosmin NO38
5	43	45.7	299	2 T25435	hypothetical prote
6	43	45.7	430	2 T39727	mannopine biosynth
7	43	45.7	683	2 S01433	repressor protein
8	42.5	45.2	264	2 C25486	K-kininogen, HMW P
9	42.5	45.2	290	2 C27115	K-kininogen, LMW P
10	42.5	45.2	315	2 A27115	major acute phase
11	42.5	44.7	639	2 A25486	kininogen, HMW I P
12	42	44.7	621	1 KGR0H1	kininogen, HMW I P
13	41	43.6	335	2 T48995	hypothetical prote
14	41	43.6	356	2 F82277	citrate (pro-3S)-1
15	41	43.6	386	2 T06484	aspartate carboxy
16	41	43.6	391	2 T06425	aspartate carboxy
17	41	43.6	457	2 S39079	puff C-8 protein -
18	41	43.6	598	2 T28238	ORF MSY077 hypote
19	41	43.6	622	2 S35122	site-specific DNA-
20	41	43.6	647	2 S38225	skeletal muscle ab
21	41	43.6	757	2 S68142	probable transcrip
22	41	43.6	1477	2 S64616	YOR1 protein - yea
23	41	43.6	1878	2 E86189	hypothetical prote
24	40	42.6	87	2 T51449	hypothetical prote
25	40	42.6	174	2 G83712	hypothetical prote
26	40	42.6	245	2 T23844	hypothetical prote
27	40	42.6	317	2 JEO175	freezed protein-1
28	40	42.6	385	2 T10181	aspartate carboxy
29	40	42.6	393	2 S60465	dom-3 protein - Ca

30	40	42.6	407	2 T30469	hypothetical prote
31	40	42.6	489	2 T36100	probable ATP-bindi
32	40	42.6	536	2 T27668	hypothetical prote
33	40	42.6	878	2 F64425	valine--tRNA ligas
34	40	42.6	1025	2 S69790	fibronectin-bindin
35	40	42.6	1035	2 T13962	sodium bicarbonate
36	40	42.6	1035	2 T14110	sodium bicarbonate
37	40	42.6	1061	2 S37667	trac-1 protein - E
38	40	42.6	1079	2 PC7034	Na+ bicarbonate co
39	40	42.6	1079	2 T14031	sodium bicarbonate
40	40	42.6	1448	2 T08526	DNA primase trac2
41	40	42.6	1448	2 S37669	trac-2 protein - E
42	40	42.6	1733	2 S27939	tensin - chicken
43	40	42.6	1744	2 A54970	tensin, cardiac mu
44	40	42.6	1785	2 S53976	probable membrane
45	40	42.6	1792	2 A57075	tensin - chicken (

ALIGNMENTS

RESULT 1

KGHUIH1
kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular
C:Species: Homo sapiens (man)
C:date: 28-May-1986 #sequence-revision 28-May-1986 #text-change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030;
R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shikawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A:title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:85122621
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:M1437; NID:9177889
R:Taekagaki, Y.; Kitamura, N.; Nakashiki, S.
J. Biol. Chem. 260, 8601-8609, 1985
A:title: Cloning and sequence analysis of cDNAs for human high molecular weight and 1
A:Reference number: A92544; MUID:85234582
A:Accession: A25276
A:Molecule type: mRNA
A:Residues: 1-592, 'I', 594-644 <TAK>
A:Cross-references: GB:M1437; NID:9186751; PIDN:AB59550.1; PID:9386852
R:Auerswald, E.A.; Roessler, D.; Mentele, R.; Asztalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A:title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:93223854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 'ANSM', 253-377 <AUE>
A:Note: differences are due to known cloning artifacts
R:Lotspesch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:title: The amino acid sequence of the light chain of human high-molecular-mass kini
A:Reference number: A91153; MUID:86030270
A:Accession: A91153
A:Molecule type: protein
A:Residues: 379-644 <LOT>
A:Note: the bradykinin sequence preceding the light chain sequence was not determined
R:Kellermann, J.; Lotspesch, F.; Henschen, A.; Muller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A:title: Completion of the primary structure of human high-molecular-mass kininogen.
A:Reference number: A24871; MUID:86108361
A:Accession: A24871
A:Molecule type: protein
A:Residues: 'Z', 20-380 <KEU>
R:Kellermann, J.; Lotspesch, F.; Henschen, A.; Muller-Esterl, W.
In Kinins IV, Greenbaum, L.W., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
A:title: Amino acid sequence of the light chain of human high molecular mass kininoge
A:Reference number: A27899
A:Accession: A27899

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <REL2>
 R:Minidrou, T.; Carretero, O.A.; Proud, D.; Malz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88209021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A31905; MUID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MA>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plasma
 A:Reference number: A34030; MUID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (Cpis) in inflammatory and
 A:Reference number: S02482; MUID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19,189-192,310-314,381-389 <LEN1>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55239; MUID:95251553
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S66059; MUID:96033974
 A:Accession: S66059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evoluti
 A:Reference number: A92545; MUID:85234583
 A:Contents: annotation; gene organization

R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:9025622
 A:Contents: annotation; bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is l
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 xproline residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379, 390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <RCH>
 F:19-131/Domain: cystatin homology <CV1>
 F:142-253/Domain: cystatin homology <CV2>
 F:264-375/Domain: cystatin homology <CV3>
 F:380-389/Product: lysyl-bradykinin (kallidin I) #status experimental <KBDY>
 F:381-389/Product: bradykinin (kallidin II) #status experimental <KBDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:28-614, 83-94, 107-126, 142-218, 229-248, 264-267, 328-340, 351-370/Disulfide bond
 F:48/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:169, 205, 294/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 593, 628/Binding site: carbohydrate (Thr) (covalent) #status
 F:577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 94; DB 1; Length 644;
 Best Local Similarity 100.0%; Pred. No. 1,5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GHRKLDDEHOGH 16
 DB 474 GHRKLDDEHOGH 489
 RESULT 2
 H64231
 protein L homolog - Mycoplasma genitalium
 C:Species: Mycoplasma genitalium
 C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 07-Dec-1999
 C:Accession: H64231
 R:Fraser, C.M.; Goeayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.
 M.; Fuhmann, J.; Nguyen, D.; Uitterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,
 C.A.; Venter, J.C.
 Science 270, 397-403, 1995
 A:Title: The minimal gene complement of Mycoplasma genitalium.
 A:Reference number: A64200; MUID:96026346
 A:Accession: H64231
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-414 <TIG>
 A:Cross-references: GB:U039709; GB:I43967; NID:g1045984; PID:g1045985; TIGR:MG288
 A:Experimental source: strain G-37
 C:Genetics:
 A:Genetic code: SGC3
 C:Superfamily: hypothetical protein MG096

Query Match 50.0%; Score 47; DB 2; Length 414;

Mol. Gen. Genet. 213, 269-277, 1988
 A:Title: The repressor gene (C) of the Streptomyces temperate phage phi-C31: nucleotide
 A:Reference number: S01433; MUID:89039715
 A:Accession number: S01433
 A:Molecule type: DNA
 A:Residues: 1-683 <SIN>
 A:Cross-references: EMBL:X12865; NID:g15458; PIDN:CAA31345.1; PID:g15459
 R:Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Charter, K.F.
 submitted to the EMBL Data Library, November 1993
 A:Reference number: S38912
 A:Accession number: S38912
 A:Molecule type: DNA
 A:Residues: 1-683 <HAR>
 A:Cross-references: EMBL:X76288; NID:g432610; PIDN:CAA53911.1; PID:g432611
 C:Genetics:
 A:Gene: c
 C:Superfamily: phage phi-C31 repressor protein C
 C:Keywords: DNA binding; transcription regulation

Query Match	45.7%	Score 43	DB 2	Length 683
Best Local Similarity	70.0%	Pred NC	40	
Matches	7	Conservative	1	Mismatches 2; Indels 0; Gaps 0
OY	7	DDDLHHGGH	16	
		:		
Db	474	DDDDVRRGGAH	483	

RESULT 8
C25486
K:kininogen, HMW precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 30-Sep-1993
C:Accession: C25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family
A:Reference number: A92625; MUID:87137443
A:Accession: C25486
A:Molecule type: DNA
A:Residues: 1-264 <RIT>
C:Comment: The nucleotide sequence was obtained from GenBank, release 55.0.
C:Superfamily: kininogen; cysstatin homology

```

Query Match 45.2%; Score 42.5; DB 2; Length 264;
Best Local Similarity 40.0%; Pred. NO. 17;
Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;

QY 1 GHKFKLKD-----DLEHOGGH 16
      ||| |||
      ||| |||
Db 86 GHQLKLDLKKOOREDDGYDHRHPVGH 110

RESULT 9
C27115
K-kininogen, LMW precursor - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text-change 20-Aug-1999
C:Accession: C27115; A25488
J:Funng, W.P.; Schreiber, G.
J: Biol. Chem. 262, 9298-9308, 1987
A:Title: Structure and expression of the genes for major acute phase alpha-1-protein (th
A:Reference number: A92653; MUID:87250580
A:Accession: C27115
A:Molecule type: DNA
A:Residues: 1-290 <FUN>
R:Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
J: Biol. Chem. 262, 2345-2351, 1987
A:Title: Differing utilization of homologous transcription initiation sites of rat K and
A:Reference number: A25488; MUID:87137465
A:Accession: A25488

```

A:Status: preliminary
A:Molecule_type: DNA
A:Residues: 1-48 <KAC>
A:Cross-references: GB:J02662; MID:g205071; PIDN:AAA41483.1; PID:g205072
A:Superfamily: Klninogen; cystatin homology
F19-65/Domain: cystatin homology (fragment) <CYS>

Query Match	45.2%	Score 42.5	DB 2	length 290
Best Local Similarity	40.0%	Pred. No. 18		
Matches 10	Conservative 1	Mismatches 5	Indels 9	Gaps 1
QY	1	GHKFKLDD-----DLEHOGGH	16	
Db	135	GHKFKLDDKQQRDDYNYRHPMGH	150	-

```

RESULT 10
A:Z7115
Major acute phase alpha-1 protein 1 - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A27115
R:Fung, W.P.; Schreiber, G.
J. Biol. Chem. 262, 9298-9308, 1987
A:Title: Structure and expression of the genes for major acute phase alpha-1-protein
A:Reference number: A92653; MUID: 87250580
A:Accession: A27115
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-315 <FUN>
C:Genetics:
A:Gene: MAP1
C:Superfamily: kininogen; cystatin homology
P:19-65/Domain: cystatin homology (fragment) <CYS>

```

Query Match	Similarity	45.2%	Score 42.5	DB 2,	length 315;
Best local	Similarity	40.0%	Pred. No. 20;		
Matches	10; Conservative	1;	Mismatches	5;	Indels 9; Gaps 1.
OY	1 GHKRLDD-----DLRHGGH	16			
Db	159 GHKRLDLDKQRRDDGYNRRPMGH	183			

RESULT 11
A25486
kininogen, HMW I precursor - rat
N:Contains: bradykinin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C:Accession: A25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family
A:Reference number: A92625; MUID:87137443
A:Accession: A25486
A:Molecule type: mRNA
A:Residues: 1-639 <KIT>
A>Note: the authors translated the codon CAA for residue 347 as Asn
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-639/Product: kininogen, HMW I #status predicted <MAT>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>

Query Match	45.28;	Score 42.5;	DB 2;	length 639;
Best Local Similarity	40.08;	Pred. No. 45;		
Matches 10;	Conservative 1;	Mismatches 5;	Indels 9;	Gaps 1;

1 GHKFLD-----DLEHGGH 16
 11: 1111 1 11 11
 Db 461 GHQLKLDLKKQREDGDYHHPVGH 485
 RESULT 12
 KGB0H1
 kininogen, HMW I precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1993 #sequence,revision 14-Nov-1993 #text,change 22-Jun-1999
 C:Accession: A01281; A91923; A91938; A29559
 R:Kilmarra, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1993
 A:Title: A single gene for bovine high molecular weight and low molecular weight kininogen
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01281
 A:Molecule type: mRNA
 A:Residues: 1-621 <KIT>
 A:Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 378-393 <KAT>
 R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Am
 A:Reference number: A91938; MUID:75170265
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 458-498 <HAN>
 R:Seiyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
 J. Biol. Chem. 262, 2768-2779, 1987
 A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
 A:Reference number: A92677; MUID:8137530
 A:Accession: A29559
 A:Molecule type: protein
 A:Residues: 72, 20-123, 125-127, 129-378 <SDE>
 R:Loitspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A:Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen
 A:Reference number: A91153; MUID:86030270
 A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R:Seiyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 808, 1984
 A:Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation; disulfide bonds
 A:Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
 C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
 C:Comment: Bradykinin, released from kininogen prior to the release of bradykinin.
 C:Comment: residue is present in the kininogen by kallikrein.
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-621/Product: HMW prokininogen I #status predicted <MNT>
 F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
 F:19-310/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:263-374/Domain: cystatin homology <CY3>
 F:379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KB DY>
 F:380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
 F:417-488/Region: glycine/histidine/lysine-rich
 F:19/Modified site: pyrroline carboxylic acid (Gln) (in mature form) #status experimen
 F:27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:

	F:167,168,169,204/Binding site:	carbohydrate (Asn)	(covalent) #status experimental
	F:136/Binding site:	carbohydrate (Thr)	(covalent) #status experimental
	F:197/Binding site:	carbohydrate (Asp)	(covalent) #status experimental
	E:378-379/Cleavage site:	Met-Lys (kallikrein)	#status experimental
	E:382/Modified site:	4-hydroxyproline (Pro)	#status predicted
	F:388-389/Cleavage site:	Arg-Ser (kallikrein)	#status experimental
	F:396,406,512/Binding site:	carbohydrate (Ser)	(covalent) #status experimental
	F:399,400,520,524,536,548,553,570/Binding site:	carbohydrate (Thr)	(covalent) #status experimental
	F:498-499/Cleavage site:	Arg-Thr (kallikrein)	#status experimental
	Query Match	44.7%;	Score 42; DB 1; Length 621;
	Best Local Similarity	50.0%;	Pred. No. 52;
	Matches	8; Conservative	0; Mismatches 8; Indels 0; Gaps 0;
Oy	1 GHKFKLDDLEHGSH	16	
DB	444 GHKNHNDGHHNGSH	459	

RESULT 13
T48995
hypothetical protein F25L23.130 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T48995
R:D/Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzini, M.; Valle, G.; Mewes, H.W.; Rudd, S.
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225012
A:Accession: T48995
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-335 <DAN>
A:Cross-references: EMBL:AL356014; GSPDB:GN00061; ATSP:F25L23.130
A:Experimental source: cultivar Columbia; BAC clone F25L23
C:Genetics:
A:Gene: ATSP:F25L23.130
A:Map position: 3
A:Introns: 115/2; 179/3; 219/3

Query Match	43.6%	Score 41	DB 2	Length 335
Best Local Similarity	50.0%	Pred. No. 38		
Matches	6	Conservative	3	Mismatches 3
				Indels 0
				Gaps 0
QY	3	KFKLDDDLHEHG	14	
				:::
db	142	KFSADDPIDHEG	153	

RESULT 14
F82277
citrinate (pro-3S)-lyase ligase VC0796 [imported] - Vibrio cholerae (strain N16961 sero
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: F82277
R;Heldelberg, J.F.; Elsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.
A.; Charlson, D.; Ermolava, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833
A;Accession: F82277
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <HEI>
A;Cross-references: GB:AE004166; GB:AE003852; NID:96655259; PIDN:AAF93960.1; GSPDB:GN
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0796
A;Map position: 1
A;Superfamily: citrinate (pro-3S)-lyase ligase

Query Match 43.6%; Score 41; DB 2; Length 356;
 Best Local Similarity 54.5%; Pred. No. 41;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 HKFKLDDLEH 12
 I : : : : :
 DB 33 HQLTVDDVEH 43

RESULT 15

T06484
 aspartate carbamoyltransferase (EC 2.1.3.2) - garden pea
 C:Species: Pisum sativum (garden pea)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 29-Sep-1999
 C:Accession: T06484
 R:Williamson, C.L.; Slocum, R.D.
 Plant Physiol. 105, 377-384, 1994
 A:Title: Molecular cloning and characterization of the PYRB1 and PYRB2 genes encoding as
 A:Reference number: Z15710; M0ID:94302176
 A:Accession: T06484
 A>Status: Preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-386 <WIL>
 A:Cross-references: EMBL:M06961; NID:g495656; PIDN:AAA62443.1; PID:g495659
 A:Experimental source: cv. Wando, leaf
 C:Genetics:
 A:Gene: PYRB1
 C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase
 C:Keywords: pyrimidine nucleotide biosynthesis; transferase

Query Match 43.6%; Score 41; DB 2; Length 386;
 Best Local Similarity 61.5%; Pred. No. 45;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHKFKLDDLEH 13
 I : : : : :
 DB 75 GOKFOIDDVIEAQ 87

Search completed: July 6, 2001, 09:18:02
 Job time: 648 sec

; MOLECULE TYPE: peptide
US-08-716-873-24

Query Match 43.6%; Score 41; DB 4; Length 1477;

Best Local Similarity 46.7%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKFKLDDLEHOGH 16
11111 :1111:
DB 1352 HKFKLDDLEHOGH 1366

RESULT 13
US-09-136-652-2
; Sequence 2, Application US/09136652A

; Patent No. 6096517
; GENERAL INFORMATION:

; APPLICANT: BRILL, ANTOINE
; APPLICANT: KHANDOUDI, NASSIRAH

; APPLICANT: MARTIN, XAVIER
; APPLICANT: BORON, WALTER

; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30010

; CURRENT APPLICATION NUMBER: US/09/136,652A
; CURRENT FILING DATE: 1998-08-19

; EARLIER APPLICATION NUMBER: EP 974019473.3
; EARLIER FILING DATE: 1997-08-19

; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 3.0

; SEQ ID NO: 2
; LENGTH: 1079

; TYPE: PRT
; ORGANISM: HOMO SAPIENS

US-09-136-652-2

Query Match 42.6%; Score 40; DB 3; Length 1079;

Best Local Similarity 37.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GHFKLDDLEHOGH 16
1 :11111
DB 410 GENVQMGCDTPHDGH 425

RESULT 14

US-09-060-410-2
; Sequence 2, Application US/09060410

; Patent No. 6165461
; GENERAL INFORMATION:

; APPLICANT: Cobb, Melanie
; APPLICANT: Hutchinson, Michele

; APPLICANT: Chen, Zhu
; APPLICANT: Berman, Kevin

; TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE
; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle

; STATE: Washington
; COUNTRY: USA

; ZIP: 98104
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,410

; FILING DATE: 14-APR-1998

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 860098.421
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1001 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-060-410-2

Query Match 41.5%; Score 39; DB 4; Length 1001;

Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 HKFKLDDLEHQ 13
1 :1111111
DB 492 HRLRLDKDLEHQ 503

RESULT 15

US-08-679-635A-4
; Sequence 4, Application US/08679635A

; Patent No. 5985643
; GENERAL INFORMATION:

; APPLICANT: Tomasz, Alexander
; APPLICANT: Delencastre, Herminia

; TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF
; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.

; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; FLOOR

; CITY: Hackensack
; STATE: New Jersey

; COUNTRY: USA
; ZIP: 07601

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/679,635A
; FILING DATE: 10-JUL-1996

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-141
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 463 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO

US-08-679-635A-4

Query Match 40.4%; Score 38; DB 2; Length 463;

Best Local Similarity 69.2%; Pred. No. 1e+02;

GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: A GENE CODING FOR A PROTEIN REGULATING
TITLE OF INVENTION: AUROBASIDIN SENSITIVITY
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/423,752
FILING DATE: April 18, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/243,403
FILING DATE: May 16, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ. ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1477
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-423-752-10

Query Match 43.6%; Score 41; DB 3; Length 1477;
Best local Similarity 46.7%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKFKLDDLEHOGGH 16
DB 1352 HKFHLDOAVEEGSN 1366

RESULT 11
US-08-945-994-3
Sequence 3, Application US/08/945994
Patent No. 6043051
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: PROMOTER
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., #800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,994

FILING DATE: No. 6043051ember 6, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX:
TELEX:
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1477
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-945-994-3

Query Match 43.6%; Score 41; DB 3; Length 1477;
Best local Similarity 46.7%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKFKLDDLEHOGGH 16
DB 1352 HKFHLDOAVEEGSN 1366

RESULT 12
US-08-716-873-24
Sequence 24, Application US/08/716873
Patent No. 6194166
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: GENE REGULATING AUROBASIDIN SENSITIVITY
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,873
FILING DATE: September 20, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ. ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1477 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:26:39 ; Search time 37.59 Seconds
(without alignments)
14.581 Million cell updates/sec

Title: US-09-437-912-7

Sequence: 1 GHRKLDDEHOGH 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	47	50.0	414	1 Y288_MYCGE	P47530 mycoplasma
3	43	45.7	430	1 MS12_AGRRH	P50201 agrobacteri
4	43	45.7	683	1 RPC_BPPHC	P08979 bacterioph
5	42.5	45.2	639	1 KNG_RAT	P08934 ratius norv
6	42	44.7	196	1 CRB8_CHICK	P55164 gallus gall
7	42	44.7	621	1 KNL_BOVIN	P01044 bos taurus
8	41.5	44.1	661	1 KNG_MOUSE	O08577 mus musculu
9	41	43.6	213	1 VEA_HPV65	O07873 mus musculu
10	41	43.6	386	1 PYB1_PEA	O43086 pisum sativ
11	41	43.6	391	1 PYB3_PEA	O43064 pisum sativ
12	41	43.6	622	1 MTL1_LACLA	P35516 lactococcus
13	41	43.6	1477	1 YOR1_YEAST	P53049 saccharomyc
14	40.5	43.1	551	1 CBX4_MOUSE	O55187 mus musculu
15	40.5	43.1	558	1 CBX4_HUMAN	O00257 homo sapien
16	40	42.6	385	1 PYB2_PEA	O43087 pisum sativ
17	40	42.6	393	1 DOM3_CAEEL	O10660 caenorhabdi
18	40	42.6	624	1 STS_MOUSE	P50427 mus musculu
19	40	42.6	878	1 SYV_METJA	O58413 methanococc
20	40	42.6	1061	1 TRC4_ECOLI	P27189 escherichia
21	40	42.6	1448	1 TRCS_ECOLI	P27189 escherichia
22	40	42.6	1744	1 TENS_CHICK	O04305 gallus gall
23	40	42.6	1785	1 GLS3_YEAST	O04352 saccharomyc
24	39	41.5	436	1 CUS1_YEAST	O02354 saccharomyc
25	39	41.5	538	1 CP5D_CANNA	P16141 candida mal
26	39	41.5	570	1 DLD_ECOLI	P06149 escherichia
27	39	41.5	690	1 LIP_STAUI	P10335 staphylococ
28	39	41.5	744	1 REL1_ECOLI	P11585 escherichia
29	39	41.5	1102	1 YE20_METJA	O58815 methanococc
30	38.5	41.0	924	1 YB53_YEAST	P34308 saccharomyc
31	38	40.4	25	1 NEOU_CHICK	P34963 gallus gall
32	38	40.4	113	1 YKGI_YEAST	P35727 saccharomyc
33	38	40.4	123	1 LCA_CAMDR	P00710 camelus drc

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	644 AA.
ID	KNG_HUMAN			
AC	P01042; P01043;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].			
GN	KNG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).			
RC	TISSUE=Liver;			
RC	MEDLINE=85234582; PubMed=2989293;			
RA	Takagaki Y., Kitamura N., Nakanishi S.;			
RT	"Cloning and low molecular weight analysis of cDNAs for human high molecular weight and low molecular weight prekallinogens. Primary structures of two human prekallinogens."			
RT	J. Biol. Chem. 260:8601-8609(1985).			
RN	[2]			
RP	GENE STRUCTURE.			
RC	MEDLINE=85234583; PubMed=2989294;			
RA	Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T., Nakanishi S.;			
RT	"Structural organization of the human kininogen gene and a model for its evolution."			
RT	J. Biol. Chem. 260:8610-8617(1985).			
RN	[3]			
RP	SEQUENCE OF 1-401 FROM N.A.			
RC	MEDLINE=85122621; PubMed=6441591;			
RA	Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;			
RT	"Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kininogen."			
RT	Biochemistry 23:5691-5697(1984).			
RN	[4]			
RP	SEQUENCE OF 379-644.			
RC	MEDLINE=86030270; PubMed=4054110;			
RA	Lotspeich F., Kellermann J., Henschen A., Foertsch B., Mueller-Esterl W.;			
RT	"The amino acid sequence of the light chain of human high-molecular-mass kininogen."			
RT	Eur. J. Biochem. 152:307-314(1985).			
RN	[5]			
RP	SEQUENCE OF 381-389.			
RC	MEDLINE=90255622; PubMed=4952632;			
RA	Pierce J.V.;			
RT	"Structural features of plasma kinins and kininogens."			
RT	Fed. Proc. 27:52-57(1968).			
RN	[6]			
RP	DISULFIDE BONDS.			
RA	Sueyoshi T., Miyata T., Kato H., Iwanaga S.;			
RT	"Disulfide bonds in bovine HMW kininogens."			

O9x286 thermotoga
P52744 homo sapien
O58407 methanococc
P09291 varicella-z
O59283 corynebacte
P49730 nicotiana t
P78790 schizosacch
P06187 salmoneila
P15590 zea mays (m
P08236 homo sapien
P10787 panulirus l
O04446 homo sapien

RL Setkagaku 56:808-808(1984).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS; (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: K02566; AAA3497.1; -
 DR EMBL: M11437; AAB59550.1; -
 DR EMBL: M11438; AAB59550.1; JOINED.
 DR EMBL: M11521; AAB59550.1; JOINED.
 DR EMBL: M11522; AAB59550.1; JOINED.
 DR EMBL: M11523; AAB59550.1; JOINED.
 DR EMBL: M11524; AAB59550.1; JOINED.
 DR EMBL: M11525; AAB59550.1; JOINED.
 DR EMBL: M11526; AAB59550.1; JOINED.
 DR EMBL: M11527; AAB59550.1; JOINED.
 DR EMBL: M11528; AAB59550.1; JOINED.
 DR EMBL: M11437; AAB59551.1; -
 DR EMBL: M11438; AAB59551.1; JOINED.
 DR EMBL: M11521; AAB59551.1; JOINED.
 DR EMBL: M11522; AAB59551.1; JOINED.
 DR EMBL: M11523; AAB59551.1; JOINED.
 DR EMBL: M11524; AAB59551.1; JOINED.
 DR EMBL: M11525; AAB59551.1; JOINED.
 DR EMBL: M11526; AAB59551.1; JOINED.
 DR EMBL: M11527; AAB59551.1; JOINED.
 DR EMBL: M11528; AAB59551.1; JOINED.
 DR PIR: A01279; KGHUHL.
 DR PIR: A25276; A25276.
 DR PIR: A01280; KGHUHL.
 DR PIR: B25276; B25276.
 DR PIR: S02482; S02482.
 DR SWISS-2DPAGE: P01043; HUMAN.
 DR MIM: 228960; -
 DR InterPro: IPR000010; -
 DR InterPro: IPR002395; -
 DR Pfam: PF00031; cystatin; 3.
 DR PRINTS: PRO0334; KININOGEN.
 DR PROSITE: PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 644 KININOGEN.
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
 FT PEPTIDE 381 389 BRADYKININ.
 FT RL

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD. RES	19	19	PYRROLIDONE CARBOXYLIC ACID. INTERCHAIN.
FT	DISULFID	28	614	
FT	DISULFID	83	94	
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	205	205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	294	294	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPPLIC	402	427	
FT	VARSPPLIC	428	644	VSPHPTSMAPAOEEDSGEKGHTR -> SHLRSEYKGR
FT	CONFLICT	593	593	PPKAGAPASEREVS (IN ISOFORM LMW).
FT	SEQUENCE	644 AA; 71945 MW; 3132B4CBARFB7E CRC64;		MISSING (IN ISOFORM LMW).

Query Match 100.0%; Score 94; DB 1; Length 644;
 Best Local Similarity 100.0%; Pred. No. 6; le-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHKFKLDDLEHOGH 16
 Db 474 GHKFKLDDLEHOGH 489
 |||||

RESULT 2
 ID Y288_MYCGE STANDARD; PRT; 414 AA.
 AC P47530;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HYPOTHETICAL PROTEIN MG288.
 OS MG288.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmales; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small R.V., Sandusky M., Fuhrmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lueder T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 Science 270:397-403(1995).

CC -1- SIMILARITY: BELONGS TO THE MG032 / MG096 / MG288 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39709; -; NOT_ANNOTATED_CDS.
DR TIGR: MG288; -
KW Hypothetical protein.
SQ SEQUENCE 414 AA: 48434 MW: 6D7480CE59D8273C CRC64;

Query Match 50.0%; Score 47; DB 1; Length 414;
Best Local Similarity 46.7%; Pred. No. 2.1;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 2 HKFKDDDLHHOGGH 16
:||||:|::|:||||:
DB 291 YKFKFEIDIKYGGY 305

RESULT 3
MS12_AGRH STANDARD; PRT; 430 AA.
ID MS12_AGRH
AC P50201;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE AGROPINE SYNTHESIS REDUCTASE (EC 1.-.-.-).
GN MAS1.
OS Agrobacterium rhizogenes.
OC Plasmid pRI8196.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=359;
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE=91352070; PubMed=1909028;
RA Hansen G., Larribe M., Vaubert D., Tempe J., Biemann B.J.,
RA Montoya A.L., Chilton M.D., Brevet J.;
RT "Agrobacterium rhizogenes pRI8196 T-DNA: mapping and DNA sequence of
RT functions involved in mannopine synthesis and hairy root
RT differentiation".
RT Proc. Natl. Acad. Sci. U.S.A. 88:7763-7767(1991).
CC -1- FUNCTION: REDUCES DEOXY-FRUCTOSYL-GLUTAMINE TO MANNOPINE.
CC -1- PATHWAY: AGROPINE / MANNOPINE SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60490; AAA22101.1; -
DR HSP: P14061; 1RPV.
DR InterPro: IPR002198; -
DR Pfam: PF00106; adh_short; 1.
DR PROSITE: PS00061; ADH_SHORT; 1.
KW Plasmid; Oxidoreductase.
FT NP_BIND 203 227 NAD OR NADP (BY SIMILARITY).
FT ACT_SITE 346 346 BY SIMILARITY.
GN KNG.
SQ SEQUENCE 430 AA: 47583 MW: DD1B7EBD92834D2C CRC64;

Query Match 45.7%; Score 43; DB 1; Length 430;

Best Local Similarity 42.1%; Pred. No. 9.8;
Matches 8; Conservative 5; Mismatches 0; Indels 6; Gaps 1;

QY 4 FRIDDL-----HHOGGH 16
:|||||:|:|:|:
DB 99 FEIDDLKERDRGKHEGGY 117

RESULT 4
RPC_BPPHC STANDARD; PRT; 683 AA.
ID RPC_BPPHC
AC P08979;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE REPRESSOR PROTEIN C.
GN C.
OS Bacteriophage phi C31.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10719;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C+-NORWICH;
RX MEDLINE=89039715; PubMed=3185504;
RA Sinclair R.B., Bibb M.J.;
RT "The repressor gene (c) of the Streptomyces temperate phage phi C31:
RT nucleotide sequence, analysis and functional cloning".
RT Mol. Gen. Genet. 213:269-277(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94374705; PubMed=8088546;
RA Hartley N.M., Murphy G.O., Bruton C.J., Chater K.F.;
RT "Sequence of the essential early region of phi C31, a temperate phage
RT of Streptomyces spp. with unusual features in its lytic
RT development".
RT Gene 147:29-40(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X12865; CAA31345.1; -
DR EMBL: X76288; CAA53911.1; -
DR PIR: S01433; S01433
KW Transcription regulation; Repressor; DNA-binding.
SQ SEQUENCE 683 AA: 74077 MW: B02379D204F37D1B CRC64;

Query Match 45.7%; Score 43; DB 1; Length 683;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 DDDLEHOGGH 16
:|||||:|:|:|:
DB 474 DDDVEROGAH 483

RESULT 5
KNG_RAT STANDARD; PRT; 639 AA.
ID KNG_RAT
AC P08934; P08933;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
GN KNG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 (1)
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RA MEDLINE=87137443; PubMed=3029068;
 RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
 RT "Differing expression patterns and evolution of the rat kininogen
 gene family.";
 RL J. Biol. Chem. 262:2190-2198(1987).
 RN (2)
 RP SEQUENCE FROM N.A. (LMW ISOFORM).
 RA MEDLINE=86008264; PubMed=2413018;
 RA Furuto Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
 RT "Primary structures of the mRNAs encoding the rat precursors for
 bradykinin and T-kinin. Structural relationship of kininogens with
 major acute phase protein and alpha 1-cysteine proteinase
 inhibitor.";
 RL J. Biol. Chem. 260:12054-12059(1985).
 RN (3)
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN-BUFFALO;
 RX MEDLINE=87250580; PubMed=2439509;
 RA Fung W.-P., Schreiber G.;
 RT "Structure and expression of the genes for major acute phase alpha 1-
 protein (thioesterin) and kininogen in the rat.";
 RL J. Biol. Chem. 262:9298-9308(1987).
 RN (4)
 RP SEQUENCE OF 1-41 FROM N.A.
 RC STRAIN-WISTAR; TISSUE=Liver;
 RX MEDLINE=87137465; PubMed=3818598;
 RA Kageyama R., Kitamura N., Okubo H., Nakanishi S.;
 RT "Differing utilization of homologous transcription initiation sites
 of rat kappa and gamma kininogen genes under inflammation condition.";
 RL J. Biol. Chem. 262:2345-2351(1987).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS; (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
 CC HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
 CC T-II.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L29428; AAA41486.1; -
 CC EMBL: M11884; AAA41487.1; -
 CC EMBL: M14369; AAA41484.1; -
 CC EMBL: M14369; AAA41485.1; ALT_SEQ.

DR EMBL: M16455; AAA41482.1; -
 DR PIR: A25486; A25486.
 DR PIR: A28055; A28055.
 DR InterPro: IPRO00010; -
 DR InterPro: IPRO002395; -
 DR Pfam: PF00031; cystatin; 3.
 DR PRINTS: PR00334; KININOGEN.
 DR PROSITE: PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing; Multigene family.
 KM
 FT SIGNAL 1 18
 FT CHAIN 19 639 KININOGEN.
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
 FT PEPTIDE 381 389 BRADYKININ.
 FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 FT DOMAIN 439 514 HIS-RICH.
 FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 107 126 BY SIMILARITY.
 FT DISULFID 142 145 BY SIMILARITY.
 FT DISULFID 206 218 BY SIMILARITY.
 FT DISULFID 229 248 BY SIMILARITY.
 FT DISULFID 264 267 BY SIMILARITY.
 FT DISULFID 328 328 BY SIMILARITY.
 FT DISULFID 351 370 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 402 433 VSPSTIARVQERDQNGEPFHGWHLAKQ -> RLINS
 FT VARSPLIC CEYKGRLLKAGGAPAPERRQAEASTVTP (IN ISOFORM
 FT LMW).
 FT VARSPLIC 434 639 MISSING (IN ISOFORM LMW).
 FT CONFLICT 61 61 E -> K (IN REF. 2)
 FT SEQUENCE 639 AA; 70933 MW; D3172DF94FF56AF5 CRC64;
 SQ
 Query Match 45.2%; Score 42.5; DB 1; Length 639;
 Best local Similarity 40.0%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;
 QY 1 GHKFKLDD-----DLEHGGH 16
 DB 461 GHQLKLDLKKQQRDGYDRHHPVGH 485
 RESULT 6
 CRBB_CHICK
 ID CRBB_CHICK STANDARD; PRT; 196 AA.
 AC P53164;
 DT 01-OCT-1996 (rel. 34, Created)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 01-OCT-2000 (rel. 40, Last annotation update)
 DE BETA CRYSTALLIN A2 (BETA-A2-CRYSTALLIN).
 GN CRYBA2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE LEGHORN; TISSUE=Lens;
 RX MEDLINE=96296047; PubMed=8674507;
 RA Duncan M.K., Banerjee-Basu S., McDermott J.B., Piatigorsky J.;
 RT "Sequence and expression of chicken beta A2- and beta
 B3-crystallins.";

RL Exp. Eye Res. 62:111-119(1996).
 CC -1- FUNCTION: CRYSTALLINS ARE THE DOMINANT STRUCTURAL COMPONENTS
 CC OF THE VERREBRATE EYE LENS.
 CC -1- SUBUNIT: HOMO/HETERODIMER, OR COMPLEXES OF HIGHER ORDER. THE
 CC STRUCTURE OF BETA-CRYSTALLIN OLIGOMERS SEEMS TO BE STABILIZED
 CC THROUGH INTERACTIONS BETWEEN THE N-TERMINAL ARMS (BY SIMILARITY).
 CC -1- DOMAIN: HAS A TWO-DOMAIN BETA-STRUCTURE, FOLDED INTO FOUR
 CC VERY SIMILAR GREEK KEY MOTIFS.
 CC -1- SIMILARITY: BELONGS TO THE BETA/GAMMA-CRYSTALLIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U28145; AAB08774.1; -.
 CC HSSP: P02522; 1BLB.
 CC InterPro: IPR001064; -.
 CC Pfam: PF00030; crystal1.2.
 CC PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 3.
 CC Eye lens protein; Duplication.
 CC FT DOMAIN 1 11 N-TERMINAL ARM.
 CC FT DOMAIN 12 51 MOTIF 1.
 CC FT DOMAIN 52 99 MOTIF 2.
 CC FT DOMAIN 100 104 CONNECTING PEPTIDE.
 CC FT DOMAIN 105 146 MOTIF 3.
 CC FT DOMAIN 147 196 MOTIF 4.
 CC SQ SEQUENCE 196 AA; 23004 MW; C177575CA4D3EE44 CRC64;
 QY 1 GHKFLDD 9
 Db 117 GHKFLSD 125
 RESULT 7
 ID KNL_BOVIN STANDARD; PRT; 621 AA.
 AC P01044;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE KININOGEN, HMW I PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS:
 DE BRADYKININ].
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84014106; PubMed=6571699;
 RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nakanishi S.;
 RT "A single gene for bovine high molecular weight and low molecular
 RT weight kininogens.";
 RL Nature 305:545-549(1983).
 RN [2]
 RP SEQUENCE OF 19-378.
 RX MEDLINE=87137530; PubMed=3546295;
 RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
 RT Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion.";
 RL J. Biol. Chem. 262:2768-2779(1987).
 RN [3]

RP SEQUENCE OF 378-393.
 RX MEDLINE=70180420; PubMed=4986212;
 RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 RT bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [4]
 RP SEQUENCE OF 458-498.
 RX MEDLINE=75170265; PubMed=1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 RT kininogen. Amino acid sequence of a fragment ('histidine-rich
 RT peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES. (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII. (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES. (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
 CC NATURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3),
 CC (4E4) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE RNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 400.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: V01491; CAA24735.1; -.
 CC PIR: A01281; KGB0H1.
 CC PIR: A29559; A29559.
 CC InterPro: IPR000010; -.
 CC InterPro: IPR002395; -.
 CC Pfam: PF00031; cystatin; 3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSSTATIN; 2.
 CC Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 CC Thiol protease inhibitor; Bradykinin; Blood coagulation;
 CC Inflammatory response; Signal.
 CC SIGNAL 1 18
 CC FT CHAIN 1 621 KININOGEN, HMW I.
 CC FT CHAIN 19 378 HEAVY CHAIN.
 CC FT PEPTIDE 380 388 BRADYKININ.
 CC FT CHAIN 389 621 LIGHT CHAIN.
 CC FT DOMAIN 19 135 CYSSTATIN-LIKE 1.
 CC FT DOMAIN 136 257 CYSSTATIN-LIKE 2.
 CC FT DOMAIN 258 378 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 19 19 N-LINKED (GLCNAC. . .).
 CC FT CARBOHYD 87 136 PARTIAL.
 CC FT CARBOHYD 136 168 OR 169.
 CC FT CARBOHYD 168 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 CC FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .).
 CC FT CARBOHYD 204 204 INTERCHAIN.
 CC FT DISULFID 27 591
 CC FT DISULFID 82 93

```

FT DISULFID 106 125
FT DISULFID 141 144
FT DISULFID 205 217
FT DISULFID 228 247
FT DISULFID 263 266
FT DISULFID 327 339
FT DISULFID 350 369
SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;

Query Match 44.7%; Score 42; DB 1; Length 621;
Best Local Similarity 50.08; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 GHKFKLDLDLEHOGGH 16
    ||| ||| |||
Db 444 GHKFKLDLDLEHOGGH 459

RESULT 8
KNG_MOUSE STANDARD; PRT; 661 AA.
AC 008677; 008676;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
GN KNG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC STRAIN=C57BL/6 x CBA; TISSUE=Liver.
RA Takano M., Kondoh J., Yajima K., Okamoto H.;
RT "Molecular cloning of cDNAs for mouse low- and high- molecular
RT kininogen.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC CLOTTING (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: HMW (SHOWN HERE) AND LMW; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, D84435; BAA19743.1; -
CC EMBL, D84435; BAA19742.1; -

```

```

DR MGD; MG1:1097705; Kng.
DR InterPro: IPR000010; -
DR InterPro: IPR0002395; -
DR InterPro: IPR003243; -
DR Pfam: PF00031; cystatin, 3.
DR PRINTS: PR00334; KININOGEN.
DR PROSITE: PS00287; CYSTATIN; 1.
KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW Alternative splicing.
FT SIGNAL 1 18
FT CHAIN 19 661
FT CHAIN 19 379
FT CHAIN 380 388
FT CHAIN 389 661
FT DOMAIN 19 135
FT DOMAIN 136 257
FT DOMAIN 258 379
FT DOMAIN 439 524
FT DOMAIN 28 631
FT DISULFID 83 94
FT DISULFID 107 125
FT DISULFID 141 144
FT DISULFID 205 217
FT DISULFID 228 247
FT DISULFID 263 266
FT DISULFID 327 339
FT DISULFID 350 369
FT CARBOHYD 82 82
FT CARBOHYD 168 168
FT CARBOHYD 204 204
FT CARBOHYD 242 242
FT VARSPPLIC 401 432
FT FT
FT FT
FT VARSPPLIC 433 661
SQ SEQUENCE 661 AA; 73102 MW; 774460258D58796E CRC64;

Query Match 44.1%; Score 41.5; DB 1; Length 661;
Best Local Similarity 40.08; Pred. No. 28;
Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;

OY 1 GHKFKLD-----DLEHOGGH 16
    ||: ||| ||| |||
Db 471 GHKFKLDYLRHQREDDDDHTHTVGH 495

RESULT 9
VE4_HP65 STANDARD; PRT; 213 AA.
ID VE4_HP65
AC 007873;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE PROBABLE E4 PROTEIN.
OS Human papillomavirus type 65.
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
OX NCBI_TaxID=28312;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-93276568; PubMed-8389082;
RA Egawa K., Dellus H., Matsukura T., Kawashima M., de Villiers E.M.;
RT "Two novel types of human papillomavirus, HPV 63 and HPV 65;
RT comparisons of their clinical and histological features and DNA
RT sequences to other HPV types.";
RL Virology 194:789-799(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X70829; CAA50175.1; -.
CC Early protein.
KW
SQ SEQUENCE 213 AA; 24669 MW; D6DD198302D24D43 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 213;
Best Local Similarity 50.0%; Pred. No. 9.4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHKFKLDLLEHOG 14
ID 142 GYEDDDDDKRENG 155

RESULT 10

ID PYB1_PEA STANDARD; PRT; 386 AA.

AC 043086; 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ASPARTATE CARBAMOYLTRANSFERASE 1 PRECURSOR (EC 2.1.3.2) (ASPARTATE TRANS-CARBAMYLASE 1) (ATCSE 1).

OS Pisum sativum (Garden pea).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

OC Fabales; Fabaceae; Papilionoideae; Pisum.

OX NCBI_TaxID=3888;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

EMBL: M96981; AAA62443.1; -.
InterPro: IPR002029; -.
InterPro: IPR002082; -.
Pfam: PF00185; OTCase; 1.
PRINTS: PR00100; AOTCase.
DR PRINTS: PR00101; ATCase.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
KW Pyrimidine biosynthesis; Transferase; Chloroplast; Transit peptide; Multigene family.
FT TRANSIT 1
FT CHAIN 1
SQ SEQUENCE 386 AA; 42617 MW; A9440F45474E29F4 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 386;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHKFKLDLLEHOG 13
ID 75 GQKFLDDVIEAQ 87

RESULT 11

ID PYB3_PEA STANDARD; PRT; 391 AA.

AC 043064; 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ASPARTATE CARBAMOYLTRANSFERASE 3 PRECURSOR (EC 2.1.3.2) (ASPARTATE TRANS-CARBAMYLASE 3) (ATCSE 3).

OS Pisum sativum (Garden pea).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

OC Fabales; Fabaceae; Papilionoideae; Pisum.

OX NCBI_TaxID=3888;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

EMBL: U05293; AAB67857.1; -.
InterPro: IPR002029; -.
InterPro: IPR002082; -.
Pfam: PF00185; OTCase; 1.
PRINTS: PR00100; AOTCase.
DR PRINTS: PR00101; ATCase.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
KW Pyrimidine biosynthesis; Transferase; Chloroplast; Transit peptide; Multigene family.
FT TRANSIT 1
FT CHAIN 1
SQ SEQUENCE 391 AA; 44343 MW; 5693EBC634FAB9A9 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 391;
Best Local Similarity 61.5%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHKFKLDLLEHOG 13
ID 80 GQKFLDDVIEAQ 92

RESULT 12

ID MTIL_LACLA STANDARD; PRT; 622 AA.

AC P35516; 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MODIFICATION METHYLASE L1AI (EC 2.1.1.72) (ADENINE-SPECIFIC

DE METHYLTRANSFERASE (LAI) (M.LAI).
GN LLAIM.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OG plasmid pTR2030.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=91294179; PubMed=1906061;
RA Hill C., Miller L.A., Kleenhammer T.R.;
RT "in vivo genetic exchange of a functional domain from a type II A
RT methylase between Lactococcal plasmid pTR2030 and a virulent
RT bacteriophage."
RT J. Bacteriol. 173:4363-4370(1991).
CC -1- FUNCTION: METHYLATION OF SPECIFIC ADENINE RESIDUES. REQUIRED FOR
CC BOTH RESTRICTION AND MODIFICATION ACTIVITIES. MAY RECOGNIZE A
CC NONPALINDROMIC SEQUENCE.
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA ADENINE =
CC S-ADENOSYL-L-HOMOCYSTEINE + DNA 6-METHYLMINOPURINE.
CC -1- SIMILARITY: CONTAINS TWO COPIES OF A SEGMENT OF FOUR AMINO ACIDS
CC WHICH IS CHARACTERISTIC OF ADENINE-SPECIFIC METHYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U17233; AAA65073.1; -
CC DR PIR: A47029; A47029.
CC DR REBASE: 3437; M.LAI.
CC DR InterPro: IPR002052; -
CC DR InterPro: IPR002294; -
CC DR Pfam: PF02086; Methyltransferase12; 2.
CC DR PRINTS: PR00505; D12NMTFRASE.
CC DR PROSITE: PS00092; N6_MTASE; 1.
CC KM Transferrase: Methyltransferase; Restriction system: Repeat; Plasmid.
CC SO SEQUENCE 622 AA; 72512 MW; 69A817F46BP9C772 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 622;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GKKFKLDLDEHOG 14
DB 567 GKKFKMLSNVLEHKG 580

RESULT 13
YOR1_YEAST
ID YOR1_YEAST STANDARD: PRT; 1477 AA.
AC P53049;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE OLIGOMYCIN RESISTANCE ATP-DEPENDENT PERMEASE YOR1.
GN YOR1 OR YGR281M.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96069397; PubMed=8524254;
RA Katkman D.J., Hallstrom T.C., Voet M., Wysock W., Golin J.,
RA Volckaert G., Moye-Rowley W.S.;
RT *Expression of an ATP-binding cassette transporter-encoding gene
RT (YOR1) is required for oligomycin resistance in Saccharomyces
RT cerevisiae.*;

RL MOL. Cell. Biol. 15:6875-6883(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97245295; PubMed=9090054;
RA Volckaert G., Voet M., Robben J.;
RT *Sequence analysis of a near-subtletomic 35.4 kb DNA segment on the
RT right arm of chromosome VII from Saccharomyces cerevisiae carrying
RT the MAL1 locus reveals 15 complete open reading frames, including
RT ZUO1, BSL2 and B102 genes and an ABC transporter gene.*;
RL Yeast 13:251-259(1997).
CC -1- FUNCTION: REQUIRED FOR OLIGOMYCIN RESISTANCE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z73066; CAA97312.1; -
CC DR HSSP: P13569; YNB.
CC DR SGD: S0003513; YOR1.
CC DR InterPro: IPR001140; -
CC DR InterPro: IPR001617; -
CC DR Pfam: PF00664; ABC_membrane; 2.
CC DR PROSITE: PS00005; ABC_tran; 2.
CC KM ATP-binding: Transmembrane; Glycoprotein; Transport.
CC SO SEQUENCE 207
FT TRANSMEM 250
FT TRANSMEM 270
FT TRANSMEM 329
FT TRANSMEM 349
FT TRANSMEM 350
FT TRANSMEM 370
FT TRANSMEM 434
FT TRANSMEM 454
FT TRANSMEM 479
FT TRANSMEM 499
FT TRANSMEM 616
FT TRANSMEM 636
FT TRANSMEM 893
FT TRANSMEM 913
FT TRANSMEM 941
FT TRANSMEM 1028
FT TRANSMEM 1048
FT TRANSMEM 1138
FT TRANSMEM 1142
FT TRANSMEM 1162
FT NP_BIND 621
FT NP_BIND 628
FT NP_BIND 1247
FT NP_BIND 1254
FT CARBOHYD 16
FT CARBOHYD 295
FT CARBOHYD 295
FT CARBOHYD 661
FT CARBOHYD 759
FT CARBOHYD 759
FT CARBOHYD 799
FT CARBOHYD 1345
FT CARBOHYD 1345
FT CARBOHYD 1366
FT CARBOHYD 1366
SQ SEQUENCE 1477 AA; 166727 MW; 40C5D36CA9B6A8C5 CRC64;

Query Match 43.6%; Score 41; DB 1; Length 1477;
Best Local Similarity 46.7%; Pred. No. 84;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKFKLDLDEHOGH 16
DB 1352 HKFHLDQAVDEBSN 1366

RESULT 14
CBX4_MOUSE
ID CBX4_MOUSE STANDARD: PRT; 551 AA.
AC O55187;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE CHROMOX PROTEIN HOMOLOG 4 (POLYCOMB 2 HOMOLOG) (PC2) (MPC2).
 GN CBX4 OR MPC2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98035734; PubMed-9367786;
 RA Alkema M.J., Jacobs J., Voncken J.W., Jenkins N.A., Copeland N.G.,
 RA Satijn D.P.E., Otte A.P., Berns A., van Lohuizen M.;
 RT "MPC2, a new murine homolog of the Drosophila polycomb protein is a
 RT member of the mouse polycomb transcriptional repressor complex";
 RL J. Mol. Biol. 273:993-1003(1997).
 CC -1- FUNCTION: INVOLVED IN MAINTAINING THE TRANSCRIPTIONALLY REPRESSIVE
 CC STATE OF GENES. MODIFIES CHROMATIN, RENDERING IT HERITABLY CHANGED
 CC IN ITS EXPRESSIBILITY.
 CC -1- SUBUNIT: COMPONENT OF THE CHROMATIN-ASSOCIATED POLYCOMB COMPLEX
 CC (PC2).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: CONTAINS 1 'CHROMO' DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; 063387; AAB96874.1; -;
 CC DR HSSP; P23197; IAP0.
 CC DR MGD; MG1:1195985; Cbx4.
 CC DR InterPro: IPR000953; -;
 CC DR Pfam; PF00385; Chromo_1.
 CC DR PRINTS; PR00504; CHROMODOMAIN.
 CC DR PROSITE; PS00598; CHROMO_1; 1.
 CC DR PROSITE; PS50013; CHROMO_2; 1.
 CC KW Chromatin regulator; Nuclear protein; Transcription regulation;
 CC KW Repressor.
 CC FT DOMAIN 11 69 CHROMO.
 CC FT DOMAIN 383 395 POLY-HIS.
 CC FT SEQUENCE 551 AA; 60581 MW; 30CEB09A82C58400 CRC64;
 SQ

Query Match 43.1%; Score 40.5; DB 1; Length 551;
 Best Local Similarity 44.4%; Pred. No. 33;
 Matches 8; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

OY 2 HKFKLDD--DLEHOGGH 16
 DB 160 HRYQPPKMYDLYQSGH 177

RESULT 15
 CBX4 HUMAN STANDARD; PRT; 558 AA.
 AC 000257;
 DT 01-OCT-2000 (Rel. 40; Created)
 DT 01-OCT-2000 (Rel. 40; Last sequence update)
 DT 01-OCT-2000 (Rel. 40; Last annotation update)
 DE CHROMOX PROTEIN HOMOLOG 4 (POLYCOMB 2 HOMOLOG) (PC2) (HPC2).
 GN CBX4.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE-97459707; PubMed-9315667;
 RA Satijn D.P.E., Olson D.J., van der Vlag J., Hamer K.M., Lambrechts C.,
 RA Masselink H., Gunster M.J., Sewalt R.G.A.B., van Driel R., Otte A.P.;

RT "Interference with the expression of a novel human polycomb protein,
 RT hpc2, results in cellular transformation and apoptosis.";
 RT Mol. Cell. Biol. 17:6076-6086(1997).
 RN [2]
 RP SEQUENCE OF 455-558 FROM N.A.
 RX MEDLINE-97342649; PubMed-9199346;
 RA Satijn D.P.E., Gunster M.J., van der Vlag J., Hamer K.M., Schul W.,
 RA Alkema M.J., Saurin A.J., Freemont P.S., van Driel R., Otte A.P.;
 RT "RING1 is associated with the polycomb group protein complex and acts
 RT as a transcriptional repressor";
 RL Mol. Cell. Biol. 17:4105-4113(1997).
 CC -1- FUNCTION: INVOLVED IN MAINTAINING THE TRANSCRIPTIONALLY REPRESSIVE
 CC STATE OF GENES. MODIFIES CHROMATIN, RENDERING IT HERITABLY CHANGED
 CC IN ITS EXPRESSIBILITY.
 CC -1- SUBUNIT: COMPONENT OF THE CHROMATIN-ASSOCIATED POLYCOMB COMPLEX
 CC (PC2).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: UBQUITOUS.
 CC -1- SIMILARITY: CONTAINS 1 'CHROMO' DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF01956; AAB80718.1; -;
 CC DR EMBL; U94344; AAB62734.1; -;
 CC DR MIM; 603079; -;
 CC DR HSSP; P23197; IAP0.
 CC DR InterPro: IPR000953; -;
 CC DR Pfam; PF00385; Chromo_1.
 CC DR PRINTS; PR00504; CHROMODOMAIN.
 CC DR PROSITE; PS00598; CHROMO_1; 1.
 CC DR PROSITE; PS50013; CHROMO_2; 1.
 CC KW Chromatin regulator; Nuclear protein; Transcription regulation;
 CC KW Repressor.
 CC FT DOMAIN 16 69 CHROMO.
 CC FT DOMAIN 383 398 POLY-HIS.
 CC FT DOMAIN 499 508 POLY-ALA.
 CC FT SEQUENCE 558 AA; 61228 MW; 7158526991D33463 CRC64;
 SQ

Query Match 43.1%; Score 40.5; DB 1; Length 558;
 Best Local Similarity 44.4%; Pred. No. 34;
 Matches 8; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

OY 2 HKFKLDD--DLEHOGGH 16
 DB 158 HRYQPPKMYDLYQSGH 175

Search completed: July 6, 2001, 09:26:40
 Job time: 971 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:25:55 ; Search time 118.42 Seconds
(without alignments)
17.876 Million cell updates/sec

Title: US-09-437-912-7
Perfect score: 94
Sequence: 1 GHRKLDDEHOGGH 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organella:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-unclassified:*
13: SP-vertebrate:*
14: SP-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	47.9	236	10	081250
2	45	47.9	239	10	081253
3	45	47.9	239	10	09SBR1
4	44	47.9	410	4	09NXU1
5	44	46.8	279	4	09VDL1
6	44	46.8	541	2	09ZSA6
7	44	46.8	681	5	09VEA9
8	43	45.7	187	10	09FTI9
9	43	45.7	296	13	091802
10	43	45.7	299	5	022856
11	43	45.7	683	9	09T215
12	43	45.7	2747	2	09L800
13	42.5	45.2	126	11	009016
14	42	44.7	648	5	09VXP2
15	42	44.7	718	5	09XYT9
16	42	44.7	1826	5	09Y255
17	41	43.6	335	10	09LX44
18	41	43.6	356	2	09KTU2
19	41	43.6	457	5	026227

20	41	43.6	477	4	09Y577	09Y577 homo sapien
21	41	43.6	499	3	P87212	P87212 polyporacea
22	41	43.6	598	14	09YX15	09YX15 melanoplus
23	41	43.6	757	4	013355	013355 homo sapien
24	41	43.6	856	5	09V706	09V706 drosophila
25	41	43.6	920	4	043178	043178 homo sapien
26	41	43.6	959	5	018359	018359 drosophila
27	41	43.6	1164	5	09YCE7	09YCE7 drosophila
28	41	43.6	1235	4	09H0E9	09H0E9 homo sapien
29	41	43.6	1878	10	09S107	09S107 arabidopsis
30	40	42.6	87	10	09LFO1	09LFO1 arabidopsis
31	40	42.6	174	2	09KFF6	09KFF6 bacillus ha
32	40	42.6	245	5	021616	021616 caenorhabdi
33	40	42.6	262	14	09IP06	09IP06 bovine vira
34	40	42.6	262	14	09IP05	09IP05 bovine vira
35	40	42.6	268	10	09LWC2	09LWC2 oryza sativ
36	40	42.6	311	2	09KWM8	09KWM8 pseudomonas
37	40	42.6	314	11	09WU66	09WU66 mus musculu
38	40	42.6	315	6	09XSC1	09XSC1 bos taurus
39	40	42.6	317	4	014780	014780 homo sapien
40	40	42.6	407	14	09YMK8	09YMK8 lymantria d
41	40	42.6	489	2	09X873	09X873 streptomyce
42	40	42.6	536	5	045994	045994 caenorhabdi
43	40	42.6	596	11	09JUG2	09JUG2 mus musculu
44	40	42.6	619	5	09YVY5	09YVY5 drosophila
45	40	42.6	670	4	09Y6R3	09Y6R3 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	236 AA.
081250	081250	PRELIMINARY:	PRT:	236 AA.
AC	081250;			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	GLOBULIN-1 (FRAGMENT).			
OS	Zea mays subsp. mays (maize).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;			
OC	Andropogoneae; Zea.			
OX	NCBI_TaxID=4578;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hilton H., Gaut B.S.;			
RT	*Speciation and domestication in maize and its wild relatives:			
RT	evidence from the Globulin-1 gene.;			
RL	Genetics 0:0-0(1998).			
DR	EMBL; AF064213; AAC31456.1; -.			
DR	HSSP; P50477; ICAU.			
DR	Mendel; 31892; Zeama; 1188; 31892.			
DR	InterPro; IPR000901; -.			
DR	InterPro; IPR001113; -.			
DR	Pfam; PF00546; Seedstore-7s; 1.			
DR	PROSITE; P50067; CEREASE_2; UNKNOWN_1.			
FT	NON-TER 236			
SQ	SEQUENCE 236 AA; 27019 MW; 1F3D9BD92C032E05 CRC64;			

Query Match 47.9%; Score 45; DB 10; Length 236;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY	7	DDLEHOGGH 16
DB	26	DDNLHHGGH 35
RESULT	2	
081253		PRELIMINARY: PRT: 239 AA.
ID	081253	

```

AC 081253;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE GLOBULIN-1 (FRAGMENT).
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Panicoideae;
OC Andropogoneae; Zea.
NCBI_TaxID=4578;
RN [1]
RP SEQUENCE FROM N.A.
RA Hilton H., Gaut B.S.;
RT "Speciation and domestication in maize and its wild relatives:
RT evidence from the globulin-1 gene.";
RL Genetics 0:0-0(1998).
DR EMBL: AF064216; AAC31459.1; -.
DR HSSP; P50477; ICAU.
DR Mendel; 31895; Zeama; 1188; 31895.
DR InterPro; IPR000901; -.
DR Pfam; PF00546; Seedstore_7s; 1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR NON_TER 239
SQ SEQUENCE 239 AA; 27384 MW; 628924A8D7BA7773 CRC64;

```

```

Query Match 47.9%; Score 45; DB 10; Length 239;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 DDDLEHOGGH 16
||:| | ||
Db 26 DDNLHHGGH 35

```

```

RESULT 3
ID Q9SBF1 PRELIMINARY; PRT; 239 AA.
AC Q9SBF1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE GLOBULIN-1 (FRAGMENT).
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Panicoideae;
OC Andropogoneae; Zea.
NCBI_TaxID=4578;
RN [1]
RP SEQUENCE FROM N.A.
RA Hilton H., Gaut B.S.;
RT "Speciation and domestication in maize and its wild relatives:
RT evidence from the globulin-1 gene.";
RL Genetics 0:0-0(1998).
DR EMBL: AF064218; AAC31461.1; -.
DR HSSP; P50477; ICAU.
DR InterPro; IPR000901; -.
DR Pfam; PF00546; Seedstore_7s; 1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR NON_TER 239
SQ SEQUENCE 239 AA; 27499 MW; 147C4F61F65307FA CRC64;

```

```

Query Match 47.9%; Score 45; DB 10; Length 239;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 DDDLEHOGGH 16
||:| | ||
Db 26 DDNLHHGGH 35

```

```

RESULT 4
ID Q9NXU1 PRELIMINARY; PRT; 410 AA.
AC Q9NXU1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE CDNA FLJ20055 FIS, CLONE COL00943.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hiki J. T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK000062; BAA90919.1; -.
SQ SEQUENCE 410 AA; 46716 MW; 6E9C09639AC02CF3 CRC64;

```

```

Query Match 47.9%; Score 45; DB 4; Length 410;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 GHRFKDDLEHOGGH 16
||:| | ||
Db 184 GHRFKSDITKQDGH 199

```

```

RESULT 5
ID Q9VDL1 PRELIMINARY; PRT; 279 AA.
AC Q9VDL1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CG5412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chame C.M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Bailley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003731; AAF55780.1;
 DR FlyBase: FBgn0038806; CG5412.
 SQ SEQUENCE 279 AA; 30634 MW; 1AE724980AF18201 CRC64;

Query Match 46.8%; Score 44; DB 5; Length 279;
 Best Local Similarity 57.1%; Pred. No. 29;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 KKKLUDDLEHOGG 16
 DB 205 RFRNAEVLHSGH 218

RESULT 6
 ID 0925A6 PRELIMINARY; PRT; 541 AA.

DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE PUTATIVE LONG-CHAIN-FATTY-ACID-COA LIGASE.

GN SC265.17.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;

RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=A3(2);
 RA Oliver K., Harris D.;
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.

RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.

RN [3]
 RC SEQUENCE FROM N.A.
 RP STRAIN=A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RT MOL. Microbiol. 21:77-96(1996).
 DR EMBL: AL03478; CAB36604.1;
 DR HSSP: P08659; ILCI.

DR InterPro: IPR000873;
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Ligase

Query Match 46.8%; Score 44; DB 2; Length 541;
 SEQUENCE 541 AA; 59034 MW; CEB7374431F28CE5 CRC64;

Best Local Similarity 63.6%; Pred. No. 57;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 5 KLDDLEHOGG 15
 DB 348 RMDDEHRTG 358

RESULT 7
 ID 09VEA9 PRELIMINARY; PRT; 681 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CG7156 PROTEIN.

GN CG7156.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
 RP [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Cealniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkov A., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelia S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003721; AAF55518.1;
 DR FlyBase: FBgn0038588; CG7156.

DR InterPro: IPR001683;
 DR Pfam: PF00069; Pkinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR SMART: SM00312; PK; 1.

ATP-binding; Transferase.
 KW ATP-binding; Transferase.
 SEQUENCE 681 AA; 77072 MW; E28A66C0E8384B4 CRC64;


```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NORMICH STOCK;
RX MEDLINE-99162580; PubMed-10051617;
RA Hendrix R.W., Smith M.C.M., Burns N., Ford M.E., Hatfull G.F.;
RT "All the world's a phage.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2192-2197(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NORMICH STOCK;
RX MEDLINE-99238410; PubMed-10219087;
RA Smith M.C.M., Burns N., Wilson R.N., Gregory M.A.;
RT "The complete genome sequence of the Streptomyces temperate phage C31:
evolutionary relationships to other viruses.";
RL Nucleic Acids Res. 27:2145-2155(1999).
DR EMBL: AJ006589; CA07123.1; -
SO SEQUENCE 683 AA; 73980 MW; EC114A061EBCA4BD CRC64;

Query Match
Best Local Similarity 45.7%; Score 43; DB 9; Length 683;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 DDDLEHOGGH 16
DB 474 DDDVEROGGAH 483

RESULT 12
OYLB00
ID Q9LB00 PRELIMINARY; PRT; 2747 AA.
AC Q9LB00;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE RTX PROTEIN.
GN ASX.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33658;
RA Braun M., Frey J., Kuhnert P.;
RT "280 kDa RTX protein of Aeromonas.";
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF218037; AAF27914.1; -
DR InterPro: IPR001343; -
DR InterPro: IPR002035; -
DR InterPro: IPR002048; -
DR Pfam: PF00353; hemolysinCbind; 3.
DR PRINTS: PR00313; CABDNCRPT.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE: PS00330; HEMOLYSIN_CALCITUM; 4.
SO SEQUENCE 2747 AA; 280202 MW; 208FE380E44A5F37 CRC64;

Query Match
Best Local Similarity 45.7%; Score 43; DB 2; Length 2747;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 GHKFKDDLEHOGGH 16
DB 2663 GOKLIDSLDDHGDH 2678

RESULT 13
OY09016
ID OY09016 PRELIMINARY; PRT; 126 AA.
AC OY09016;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)

```

```

DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE K-KININOGEN (FRAGMENT).
GN KNCK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DONRYU;
RX MEDLINE-97468288; PubMed-9321484;
RA Harris E.L., Grigor M.R., Innes B.A., Harrap S.B., Kolke G.,
RA Jacob H.J.;
RT "Strain-specific deletions in exon 10 of rat K-kininogen and T1-
kininogen genes allow mapping of both genes to rat chromosome 11.";
RL Mamm. genome 8:791-792(1997).
DR EMBL: AF003623; AAC09070.1; -
DR InterPro: IPR002395; -
DR PRINTS: PR00334; KININOGEN.
DR NON_TER 1 1
FT NON_TER 1 1
SO SEQUENCE 126 AA; 14092 MW; 9CCDF8751DA49C88 CRC64;

Query Match
Best Local Similarity 45.2%; Score 42.5; DB 11; Length 126;
Matches 10; Conservative 1; Mismatches 5; Indels 9; Gaps 1;

OY 1 GHKFKID-----DLEHOGGH 16
DB 49 GHQLKIDLKQOREGDYDHRHPVGH 73

RESULT 14
OYVXP2
ID Q9VXP2 PRELIMINARY; PRT; 648 AA.
AC Q9VXP2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE RHP PROTEIN.
GN RHP OR CG8497.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Dey S., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

Job time: 991 sec

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlishina N.V., Mobarry C., Morris J., Moshirei A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003500; AAF48516.1; -.
DR FlyBase: FBgn0026374; Rhp.
DR InterPro: IPR001478; -.
DR Pfam: PF00595; PDZ; 1.
DR SMART: SM00228; PDZ; 1.
SO SEQUENCE 648 AA; 72858 MW; 85EE4EBC6947D45D CRC64;

Query Match 44.7%; Score 42; DB 5; Length 648;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 DDDLEHGGH 16
DB 489 DDDERHDGCV 498

RESULT 15
ID O9XXY9 PRELIMINARY; PRT; 718 AA.
AC O9XXY9;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE RHOPILIN.
GN RHP OR CG8497.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Edwards K.A., Kameshiro K., Yamamoto D.;
RT "Mutations in the *Drosophila* Rhophilin gene at 13E.";
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF132025; AAD31273.1; -.
DR FlyBase: FBgn0026374; Rhp.
DR InterPro: IPR000861; -.
DR InterPro: IPR001478; -.
DR Pfam: PF00595; PDZ; 1.
DR Pfam: PF02185; HRI; 1.
DR SMART: SM00228; PDZ; 1.
SO SEQUENCE 718 AA; 80826 MW; AF9DD0C57132AA31 CRC64;

Query Match 44.7%; Score 42; DB 5; Length 718;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 DDDLEHGGH 16
DB 559 DDDERHDGCV 568

Search completed: July 6, 2001, 09:25:55

```

CC or send an email to license@sb-sib.ch).
CC EMBL: D13797; BAA02954.1;
DR PIR: A45969; A45969.
KW Functional; Signal; Repeat.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 85 ANTIFUNGAL PROTEIN.
FT DOMAIN 19 73 2 X 7 AA REPEATS OF Q-H-G-H-G-G-Q.
FT REPEAT 19 25 1.
FT REPEAT 67 73 2.
SQ SEQUENCE 85 AA; 9018 MW; E881779F923FB69B CRC64;

Query Match
Best Local Similarity 49.5%; Score 48.5; DB 1; Length 85;
Matches 10; Conservative 3; Mismatches 0; Indels 5; Gaps 2;

QY 1 KHGG--HGKH--KNKG 13
   :||||| 11:1 11:1
Db 67 QHGGGQHGGHGDGKKNRG 84

RESULT 10
CH38_DROME STANDARD; PRT; 306 AA.
AC P07183; Q9W3E5;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CHORION PROTEIN S38.
OS CP38 OR S38 OR CG1121.
ON Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-87246506; PubMed-3036489;
RA Spradling A.C., de Cicco D.V., Wakimoto B.T., Levine J.F.,
RT Kallayan L.J., Cooley L.;
RT "Amplification of the X-linked Drosophila chorion gene cluster
RT requires a region upstream from the s38 chorion gene.";
RL EMBO J. 6:1045-1053(1987).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blake J.R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hosten D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

```

```

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC
CC EMBL: X05245; CAA28871.1;
DR EMBL: AE003444; AAF46383.1;
DR PIR: S08607; S08607.
DR HSSP: P04002; IATP.
DR FlyBase: FBgn0000360; CP38.
KW Chorion.
SQ SEQUENCE 306 AA; 30448 MW; 2F51C96F9F82DF83 CRC64;

Query Match
Best Local Similarity 49.0%; Score 48; DB 1; Length 306;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HGHGHGKH 9
   :||||| 1
Db 207 HGHGHGKH 214

RESULT 11
YBCZ_ECOLI STANDARD; PRT; 480 AA.
AC P77485;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROBABLE SENSOR PROTEIN YBCZ (EC 2.7.3.-).
GN YBCZ.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RP SEQUENCE FROM N.A.
RC CHUNG E., Allen E., Aparicio A., Davis K., Duncan M.,
RA Federapfel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Nameth A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
CC -i- FUNCTION: PROBABLE MEMBER OF A TWO-COMPONENT REGULATORY SYSTEM
CC YBCZ/YICA. MAY ACTIVATE YICA BY PHOSPHORYLATION.
CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (PROBABLE).

```

```

CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: X16827; CAA34727.1; -
CC PIR: S08137; S08137.
CC DictyDb: DD05011; -
CC FT DOMAIN 5 26 SER-RICH.
CC FT DOMAIN 40 94 GLY-RICH.
CC SO SEQUENCE 98 AA; 8836 MW; BD7BDF649EBID50 CRC64;

Query Match 51.0%; Score 50; DB 1; Length 98;
Best Local Similarity 61.5%; Pred. No. 0.48;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 HGHGCHKRNGK 14
Db 78 HGHGCHKRNGK 90

RESULT 8
ID GAB_DROME STANDARD; PRT; 606 AA.
GAB_DROME
AC P25123; Q26302;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GAMMA-AMINOBUTYRIC-ACID RECEPTOR BETA SUBUNIT PRECURSOR (GABA(A)
DE RECEPTOR) (CYCLODIENE RESISTANCE PROTEIN).
GN RDL.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91334435; PubMed=1651498;
RA French-Constant R.H., Rochelandet T.A.,
RA MacIntyre R.J., Roush R.T.,
RT "Molecular cloning and transformation of cycloidiene resistance in
RT drosophila: an invertebrate gamma-aminobutyric acid subtype A
RT receptor locus."
RT Proc. Natl. Acad. Sci. U.S.A. 88:7209-7213(1991).
RN [2]
RP SEQUENCE OF 70-113 FROM N.A.
RX MEDLINE=93260477; PubMed=7684073;
RA French-Constant R.H., Rochelandet T.A.,
RA "Drosophila gamma-aminobutyric acid receptor gene rdl shows extensive
RT alternative splicing."
RT J. Neurochem. 60:2323-2326(1993).
CC -1- FUNCTION: GABA, AN INHIBITORY NEUROTANSITTER, MEDIATES NEURONAL
CC INHIBITION BY BINDING TO THE GABA RECEPTOR AND OPENING AN INTEGRAL
CC CHLORIDE CHANNEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- MISCELLANEOUS: RESISTANCE IS THOUGHT TO BE DUE TO INSSENSITIVITY OF
CC THE CYCLODIENE/PICTOXININ BINDING SITE ON THE GABA(A) RECEPTOR-
CC CHLORIDE IONOPHORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-6 OR MET-12 IS THE
CC INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: M69057; AAA28556.1; -
CC EMBL: S61113; AAB26669.1; -
CC PIR: A41145; A41145.
CC FlyBase: FBgn0004244; Rdl.

```

```

DR InterPro: IPR000188; -
DR InterPro: IPR001175; -
DR InterPro: IPR002289; -
DR Pfam: PF00065; neur_chan. 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR PRINTS: PR00253; GABARECEPTR.
DR PRINTS: PR01160; GABAREBETA.
DR PROSITE: PS00236; NEUROTR ION_CHANNEL. 1.
KW postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KM transmembrane. 1
FT SIGNAL 1 44
FT CHAIN 45 606
FT FT
FT FT
FT DOMAIN 45 268
FT TRANSMEM 269 291
FT TRANSMEM 297 316
FT TRANSMEM 333 356
FT DOMAIN 357 368
FT TRANSMEM 369 590
FT CARBOHYD 58 58
FT CARBOHYD 248 248
FT CARBOHYD 253 253
FT DISULFID 185 199
FT DOMAIN 499 535
FT CONFLICT 97 97
FT CONFLICT 101 101
FT SEQUENCE 606 AA; 65748 MW; 81B9DB08E1906EF1 CRC64;

Query Match 51.0%; Score 50; DB 1; Length 606;
Best Local Similarity 77.8%; Pred. No. 2.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KHGHGCHKH 9
Db 426 KHGHGCHKH 434

RESULT 9
ID ANTF_SARPE STANDARD; PRT; 85 AA.
AC 008617;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ANTI-FUNGAL PROTEIN PRECURSOR (AFP).
OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestrioidea; Sarcophagidae; Sarcophaga.
OX NCBI_Taxid=7386;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-54.
RX TISSUE=Fat body;
RX MEDLINE=93280179; PubMed=8505329;
RA Iijima R., Kurata S.,
RA "Purification, characterization, and cDNA cloning of an antifungal
RT protein from the hemolymph of Sarcophaga peregrina (Flesh fly)
RT larvae."
RL J. Biol. Chem. 268:12055-12061(1993).
CC -1- FUNCTION: THIS PROTEIN INHIBITS THE GROWTH OF A VARIETY OF
CC FUNGAL SPECIES. THE ANTI-FUNGAL ACTIVITY OF THIS PROTEIN IS
CC ENHANCED BY THE PRESENCE OF SARCOPTOXIN IA.
CC -1- SUBUNIT: HOMODIMER.
CC -1- TISSUE SPECIFICITY: HEMOLYMPH.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

FT DOMAIN 19 136 CYSTATIN-LIKE 1.
FT DOMAIN 137 258 CYSTATIN-LIKE 2.
FT DOMAIN 259 380 CYSTATIN-LIKE 3.
FT DOMAIN 439 514 HIS-RICH.
FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 107 126 BY SIMILARITY.
FT DISULFID 142 145 BY SIMILARITY.
FT DISULFID 206 218 BY SIMILARITY.
FT DISULFID 229 248 BY SIMILARITY.
FT DISULFID 264 267 BY SIMILARITY.
FT DISULFID 328 340 BY SIMILARITY.
FT DISULFID 351 370 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 402 433 VSPSYIARVOEERDGENEGPIHGHGWLHAKO -> RLINS
CEYKGRLLKAGGAPAPERAEASYTP (IN ISOFORM
LMW).
FT VARSPLIC 434 639 MISSING (IN ISOFORM LMW).
FT CONFLICT 61 61 E -> K (IN REF. 2).
SQ SEQUENCE 639 AA; 70933 MW; D3172DF94FF56AF5 CRC64;

```

```

Query Match Score 64; DB 1; Length 639;
Best Local Similarity 70.6%; Pred. No. 0.027;
Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

```

```

OY 2 HGHGSHG--KHKNGKKKN 16
Db 495 HGHGSHGDKHTNKDKNN 511

```

```

RESULT 6
ID HKN1_MAIZE STANDARD; PRT; 359 AA.
AC P24345;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HOMEOBOX PROTEIN KNOTTED-1.
GN KN-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91172321; PubMed=1672445; Hake S.;
RA Vollbrecht E., Veit B., Sinha N., Hake S.;
RT "The developmental gene knotted-1 is a member of a maize homeobox
RT gene family."
RL Nature 350:241-243(1991).
RN [2]
RP DEVELOPMENTAL EXPRESSION.
RX STRAIN=CV. B73;
RA Smith L.G., Green B., Veit B., Hake S.;
RT "A dominant mutation in the maize homeobox gene, knotted-1, causes its
RT ectopic expression in leaf cells with altered fates."
RL Development 116:21-30(1992).
CC -1- FUNCTION: POSSIBLE TRANSCRIPTION FACTOR THAT REGULATES GENES
CC INVOLVED IN DEVELOPMENT. MUTATIONS IN KN-1 ALTER LEAF DEVELOPMENT.
CC BUT CONTINUE TO DIVIDE, FORMING KNOTS. MAY PARTICIPATE IN THE
CC SWITCH FROM INDETERMINATE TO DETERMINATE CELL FATES. PROBABLY
CC BINDS TO THE DNA SEQUENCE 5'-TGAC-3'.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN APICAL MERISTEMS OF VEGETATIVE

```

```

CC AND FLORAL STEMS AS WELL AS IN THE UNDERLYING GROUND MERISTEM.
CC SPECIFICALLY EXPRESSED IN VASCULAR BUNDLES DEVELOPING BOTH IN THE
CC LEAF AND STEM. VERY LOW LEVELS OF EXPRESSION IN LEAVES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT APICAL AND VEGETATIVE
CC MERISTEMS DURING DEVELOPMENT. DOWN-REGULATED AS LEAVES AND FLORAL
CC ORGANS ARE INITIATED.
CC -1- SIMILARITY: BELONGS TO THE TALE/KNOX FAMILY OF HOMEOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X61308; CAA43605.1; -
CC PIR; S14283; S14283.
CC TRANSFAC; T02062; -.
CC MaizeDB; 65584; -.
DR InterPro; IPR001356; -.
DR PROSITE; PS00027; HOMEOBOX_1; 1.
DR PROSITE; PS50071; HOMEOBOX_2; 1.
DR DNA-binding; Homeobox; Transcription regulation; Nuclear protein.
FT DOMAIN 22 30 POLY-HIS.
FT DOMAIN 92 95 POLY-SER.
FT DOMAIN 264 270 POLY-LYS.
FT DOMAIN 240 263 EUK DOMAIN.
FT DNA_BIND 264 326 HOMEOBOX (TALE-TYPE).
SQ SEQUENCE 359 AA; 39827 MW; 800FFD82082400FB CRC64;

```

```

Query Match Score 51; DB 1; Length 359;
Best Local Similarity 87.5%; Pred. No. 1.2;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 HGHGSHG 9
Db 15 HGHGSHG 22

```

```

RESULT 7
ID 2C_DICDI STANDARD; PRT; 98 AA.
AC P15648;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE 2C PROTEIN.
GN 2C.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AX2;
RA MEDLINE=90205618; PubMed=2157129;
RA Ranji D.P., Richards A.J., Jagger P., Bleasby A., Hames B.D.;
RT "Two cyclic AMP-regulated genes from Dictyostelium discoideum encode
RT homologous proteins."
RL Mol. Microbiol. 4:129-135(1990).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ONLY LATE IN DEVELOPMENT. ITS
CC EXPRESSION CEASIS UPON CELL DISAGGREGATION BUT IS FULLY RESTORED
CC BY EXOGENOUS CAMP.
CC -1- DOMAIN: MAY FORM AN EXTENDED COIL STRUCTURE.
CC -1- SIMILARITY: MAY FORM AN EXTENDED COIL STRUCTURE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```


RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 RT bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [4]
 RP SEQUENCE OF 458-498.
 RA MEDLINE=75170265; PubMed=1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 RT kininogen. Amino acid sequence of a fragment (histidine-rich
 RT peptide) released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
 CC NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 400.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: V01491; CAA24735.1; -
 DR PIR: A01281; KGB0H1.
 DR PIR: A29359; A29359.
 DR InterPro: IPR000010; -
 DR InterPro: IPR002395; -
 DR Pfam: PF00031; Cystatin; 3.
 DR PRINTS: PR00334; KININOGEN.
 DR PROSITE: PS00287; CYSTATIN; 2.
 DR Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 KW Thiol protease inhibitor; Bradykinin; Blood coagulation;
 KW Inflammatory response; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 621 KININOGEN, HMW I.
 FT PEPTIDE 19 378
 FT CHAIN 389 621 BRADYKININ.
 FT CHAIN 19 135 LIGHT CHAIN.
 FT DOMAIN 136 257 CYSTATIN-LIKE 1.
 FT DOMAIN 258 378 CYSTATIN-LIKE 2.
 FT MOD. RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHD 136 136 PARTIAL.
 FT CARBOHD 168 168 OR 169.
 FT CARBOHD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 FT CARBOHD 204 204 N-LINKED (GLCNAC. . .).
 FT DISULFD 27 591 INTERCHAIN.
 FT DISULFD 82 93
 FT DISULFD 106 125
 FT DISULFD 141 144

FT DISULFD 205 217
 FT DISULFD 228 247
 FT DISULFD 263 266
 FT DISULFD 327 339
 FT DISULFD 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;
 Query Match 94.9%; Score 93; DB 1; Length 621;
 Best Local Similarity 93.8%; Pred. No. 15e-06;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 KHGHGKHKRKNKKN 16
 DB 476 KHGHGKHKRKNKKN 491
 RESULT 4
 KNG_MOUSE STANDARD; PRT; 661 AA.
 AC 008677.008676;
 DT 01-OCT-2000 (rel. 40, Created)
 DT 01-OCT-2000 (rel. 40, Last sequence update)
 DT 01-OCT-2000 (rel. 40, Last annotation update)
 DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
 OS KNG.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RC STRAIN=C57BL/6 x CBA; TISSUE=Liver;
 RA Takano M., Kondoh J., Yajima K., Okamoto H.;
 RT "Molecular cloning of cDNAs for mouse low- and high- molecular
 RT kininogen.";
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D84435; BA19743.1; -
 DR EMBL: D84415; BA19742.1; -
 DR MGD: MGI:1097705; Kng.
 DR InterPro: IPR000010; -

RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion.";
 RL J. Biol. Chem. 262:2768-2779(1987).
 RP [3]
 RP SEQUENCE OF 376-391.
 RX MEDLINE-70180420; PubMed-4986212;
 RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 RT bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [4]
 RP SEQUENCE OF 387-455.
 RX MEDLINE-76260155; PubMed-956151;
 RA Han Y.N., Kato H., Iwanaga S., Suzuki T.;
 RT "Primary structure of bovine plasma high-molecular-weight kininogen.
 RT The amino acid sequence of a glycopeptide portion (fragment 1)
 RT following the C-terminus of the bradykinin moiety.";
 RL J. Biochem. 79:1201-1222(1976).
 RN [5]
 RP SEQUENCE OF 456-496.
 RX MEDLINE-75170265; PubMed-1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 RT kininogen. Amino acid sequence of a fragment ('histidine-rich
 RT peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
 CC NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AND THE RESULT OF ALTERNATE RNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 398.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: V01492; CAA24736.1; -
 CC EMBL: V01492; CAA24737.1; ALT_SEQ.
 CC PIR: A01282; KGB0H2.
 CC PIR: B25559; B29559.
 CC DR HSSP: P04129; IAFI.
 CC DR InterPro: IPR000010; -
 CC DR InterPro: IPR002395; -
 CC DR Pfam: PF00034; cystatin.3.
 CC DR PRINTS: PR00334; KININOGEN.
 CC DR PROSITE: PS00287; CYSSTATIN; 2.
 CC DR Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 CC Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
 CC Inflammatory response.
 KW

FT SIGNAL 1 18
 FT CHAIN 19 619 KININOGEN, HMW II.
 FT CHAIN 19 376 HEAVY CHAIN.
 FT PEPTIDE 378 386 BRADYKININ.
 FT CHAIN 387 619 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSSTATIN-LIKE 1.
 FT DOMAIN 136 256 CYSSTATIN-LIKE 2.
 FT DOMAIN 257 376 CYSSTATIN-LIKE 3.
 FT MOD_RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 136 136 PARTIAL.
 FT CARBOHYD 168 168 OR 169.
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 400 400 INTERCHAIN.
 FT DISULFID 27 589
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 261 264
 FT DISULFID 325 337
 FT DISULFID 348 367
 FT VARIANT 398 398
 FT VARIANT 401 401 T -> P.
 FT VARIANT 454 454 L -> V.
 FT VARIANT 454 454 H -> K.
 SQ SEQUENCE 619 AA; 68710 MW; F04320A8EB0E0DA CRC64;
 Query Match 94.9%; Score 93; DB 1; Length 619;
 Best Local Similarity 93.8%; Pred. No. 1.5e-06;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 YQ 1 KHGGHGKHKNNKKN 16
 Db 474 KHGGHGKHKNNKKN 489
 RESULT 3
 ID KNL_BOVIN STANDARD; PRT; 621 AA.
 AC P01044;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE KININOGEN, HMW I PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS:
 DE BRADYKININ].
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-84014106; PubMed-6571699;
 RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
 RT "A single gene for bovine high molecular weight and low molecular
 RT weight kininogens.";
 RL Nature 305:545-549(1983).
 RN [2]
 RP SEQUENCE OF 19-378.
 RX MEDLINE-87137530; PubMed-3546295;
 RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
 RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion.";
 RL J. Biol. Chem. 262:2768-2779(1987).
 RN [3]
 RP SEQUENCE OF 378-393.
 RX MEDLINE-70180420; PubMed-4986212;

RL Seikagaku 56:808-808(1984).

CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)

CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY

CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO

CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-

CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE

CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF

CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE

CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND

CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A

CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR

CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF

CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS

CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,

CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)

CC LHM-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LHM-

CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD

CC CLOTTING.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>

CC or send an email to license@isb-slb.ch).

CC -----

DR EMBL: K02566; AAA35497.1; -

DR EMBL: M11437; AAB59550.1; -

DR EMBL: M11438; AAB59550.1; JOINED.

DR EMBL: M11521; AAB59550.1; JOINED.

DR EMBL: M11522; AAB59550.1; JOINED.

DR EMBL: M11523; AAB59550.1; JOINED.

DR EMBL: M11524; AAB59550.1; JOINED.

DR EMBL: M11525; AAB59550.1; JOINED.

DR EMBL: M11526; AAB59550.1; JOINED.

DR EMBL: M11527; AAB59550.1; JOINED.

DR EMBL: M11528; AAB59550.1; JOINED.

DR EMBL: M11437; AAB59551.1; -

DR EMBL: M11438; AAB59551.1; JOINED.

DR EMBL: M11521; AAB59551.1; JOINED.

DR EMBL: M11522; AAB59551.1; JOINED.

DR EMBL: M11523; AAB59551.1; JOINED.

DR EMBL: M11524; AAB59551.1; JOINED.

DR EMBL: M11525; AAB59551.1; JOINED.

DR EMBL: M11526; AAB59551.1; JOINED.

DR EMBL: M11527; AAB59551.1; JOINED.

DR EMBL: M11528; AAB59551.1; JOINED.

DR PIR: A01279; KGH0H1.

DR PIR: A25276; A25276.

DR PIR: A01280; KGH0L1.

DR PIR: B25276; B25276.

DR PIR: S02482; S02482.

DR SWISS-2DPAGE; P01043; HUMAN.

DR MIM: 228960; -

DR InterPro; IPR000010; -

DR InterPro; IPR002395; -

DR Pfam; PF000031; cystatin.3.

DR PRINTS; PR00334; KININOGEN.

DR PROSITE; PS00287; CYSTATIN.2.

DR Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;

DR Bradykinin; Blood coagulation; Inflammatory response; Signal;

DR Alternative splicing.

FT SIGNAL 1 18

FT CHAIN 19 644 KININOGEN.

FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 389 BRADYKININ.

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD.RES	19	19	
FT	DISULFID	28	614	PYRROLIDONE CARBOXYLIC ACID.
FT	DISULFID	83	94	INTERCHAIN.
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	
FT	CARBOHYD	169	169	
FT	CARBOHYD	205	205	
FT	CARBOHYD	294	294	
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPLIC	402	427	
FT	VARSPLIC	428	644	VSPPTSMAPADDERDSKKEGHTR -> SHLRSCFYKGR
FT	CONFLICT	593	593	PKRGAEPASEKEVS (IN ISOFORM LHM).
FT	SEQUENCE	644 AA; 71945 MW; 3132B4CBARF8FB7E CRC64;		MISSING (IN ISOFORM LHM).
FT	SEQUENCE			T -> I (IN REF. 1).

Query Match 100.0%; Score 98; DB 1; Length 644;

Best Local Similarity 100.0%; Pred. No. 2.8e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHGHGCHKHKNKGN 16

DB 498 KHGHGCHKHKNKGN 513

RESULT 2

ID KMH2_BOVIN STANDARD; PRT; 619 AA.

AC P01045:

DT 21-JUL-1986 (rel. 01, Created)

DT 21-JUL-1986 (rel. 01, Last sequence update)

DT 01-OCT-2000 (rel. 40, Last annotation update)

DE KININOGEN, HMW II PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxId=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84014106; PubMed=6571699;

RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakaniishi S.;

RT "A single gene for bovine high molecular weight and low molecular

RT weight kininogens.";

RL Nature 305:545-549(1983).

RN [2]

RP SEQUENCE OF 19-376.

RP MEDLINE=87137530; PubMed=3546295;

RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:26:40 ; Search time 37.59 seconds

(Without alignments)
14.581 Million cell updates/sec

Title: US-09-437-912-8

Perfect score: 98

Sequence: 1 KHGHGKHKMKGN 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	93	94.9	619	1 KNG_BOVIN	P01045 bos taurus
3	93	94.9	621	1 KNG_BOVIN	P01044 bos taurus
4	70	71.4	661	1 KNG_MOUSE	P08677 mus musculu
5	64	65.3	639	1 KNG_RAT	P08934 ratus norv
6	51	52.0	359	1 KNG_MAIZE	P24345 zea mays (m
7	50	51.0	98	1 KNG_MOUSE	P07485 escherichia
8	50	51.0	606	1 KNG_MOUSE	P07485 escherichia
9	48.5	49.5	85	1 KNG_MOUSE	P07485 escherichia
10	48	49.0	306	1 KNG_MOUSE	P07485 escherichia
11	48	49.0	480	1 KNG_MOUSE	P07485 escherichia
12	47.5	48.5	503	1 KNG_MOUSE	P07485 escherichia
13	47.5	48.5	503	1 KNG_MOUSE	P07485 escherichia
14	47	48.0	97	1 KNG_MOUSE	P07485 escherichia
15	46.5	47.4	352	1 KNG_MOUSE	P07485 escherichia
16	46.5	47.4	1213	1 KNG_MOUSE	P07485 escherichia
17	45.5	46.4	439	1 KNG_MOUSE	P07485 escherichia
18	45.5	46.4	670	1 KNG_MOUSE	P07485 escherichia
19	45	45.9	286	1 KNG_MOUSE	P07485 escherichia
20	45	45.9	946	1 KNG_MOUSE	P07485 escherichia
21	45	45.9	1273	1 KNG_MOUSE	P07485 escherichia
22	44.5	45.4	286	1 KNG_MOUSE	P07485 escherichia
23	44.5	45.4	515	1 KNG_MOUSE	P07485 escherichia
24	44	44.9	1085	1 KNG_MOUSE	P07485 escherichia
25	43	43.9	197	1 KNG_MOUSE	P07485 escherichia
26	43	43.9	261	1 KNG_MOUSE	P07485 escherichia
27	43	43.9	265	1 KNG_MOUSE	P07485 escherichia
28	43	43.9	420	1 KNG_MOUSE	P07485 escherichia
29	43	43.9	423	1 KNG_MOUSE	P07485 escherichia
30	43	43.9	515	1 KNG_MOUSE	P07485 escherichia
31	43	43.9	599	1 KNG_MOUSE	P07485 escherichia
32	43	43.9	678	1 KNG_MOUSE	P07485 escherichia
33	43	43.9	2038	1 KNG_MOUSE	P07485 escherichia

ALIGNMENTS

RESULT ID	KNG_HUMAN	STANDARD	PRT	644 AA	000454 tetrahymena
AC	P01042	P01043			P50117 oryctolagus
DT	21-JUL-1986 (Rel. 01, Created)				O46248 drosophila
DT	01-FEB-1996 (Rel. 33, Last sequence update)				O46238 drosophila
DT	01-OCT-2000 (Rel. 40, Last annotation update)				P17816 hordeum vul
DE	KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].				P20025 zea mays (m
GN	KNG.				P57513 buchena ap
OS	Homo sapiens (human).				O9v244 drosophila
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				P54289 drosophila
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				P55965 drosophila
OX	NCBI_TaxID=9606;				P38041 saccharomyc
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).				
RC	TISSUE=Liver;				
RX	MEDLINE=85234582; PubMed=2989293;				
RA	Takagaki Y., Kitamura N., Nakanishi S.;				
RT	"Cloning and sequence analysis of cDNAs for human high molecular weight and low molecular weight prekallinogens. Primary structures of				
RT	two human prekallinogens.";				
RL	J. Biol. Chem. 260:8601-8609(1985).				
RN	[2]				
RP	GENE STRUCTURE.				
RX	MEDLINE=85234583; PubMed=2989294;				
RA	Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,				
RT	Nakanishi S.;				
RT	"Structural organization of the human kininogen gene and a model for				
RT	its evolution.";				
RL	J. Biol. Chem. 260:8610-8617(1985).				
RN	[3]				
RP	SEQUENCE OF 1-401 FROM N.A.				
RX	MEDLINE=85122621; PubMed=6441591;				
RA	Ohkubo I., Kurachi K., Takasawa T., Shiohara H., Sasaki M.;				
RT	"Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and				
RT	its identity with low molecular weight kininogen.";				
RL	Biochemistry 23:5691-5697(1984).				
RN	[4]				
RP	SEQUENCE OF 379-644.				
RX	MEDLINE=86030270; PubMed=4054110;				
RA	Lottspeich F., Kellermann J., Henschen A., Foerisch B.,				
RT	Wueller-Esterl W.;				
RT	"The amino acid sequence of the light chain of human high-molecular-				
RT	mass kininogen.";				
RL	Eur. J. Biochem. 152:307-314(1985).				
RN	[5]				
RP	SEQUENCE OF 381-389.				
RX	MEDLINE=90255622; PubMed=4952632;				
RA	Pierce J.V.;				
RT	"Structural features of plasma kinins and kininogens.";				
RL	Fed. Proc. 27:52-57(1968).				
RN	[6]				
RP	DISULFIDE BONDS.				
RA	Sueyoshi T., Miyata T., Kato H., Iwanaga S.;				
RT	"Disulfide bonds in bovine HMW kininogens.";				

QY 2 HGHGKHKHKKN 16
DB 32 HGHGKHKHKKN 46

RESULT 13

T26522
hypothetical protein Y18D10A.10 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C.Accession: T26522
R.Harris B
Submitted to the EMBL Data Library, December 1998
A.Reference number: Z20226
A.Accession: T26522
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 241 <YIP>
A.Cross-references: EMBL:AI034393; PIDN:CAA22313.1; CESP:Y18D10A.10
A:Experimental source: clone Y18D10A
C:Genetics:
A:Gene: CESP:Y18D10A.10
A:Introns: 99/3; 127/1; 193/2

Query Match 51.0%; Score 50; DB 2; Length 241;
Best Local Similarity 63.6%; Pred. No. 4.4;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 HGHGKHKHKKN 12
DB 47 HGHGKHKHKKN 57

RESULT 14

A11145
gamma-aminobutyric acid receptor A, cyclodiene resistance-conferring - fruit fly (Drosophila)
C.Species: Drosophila melanogaster
C.Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 21-Aug-1998
C.Accession: A11145
R.French-Constant, R.H.; Mottlock, D.P.; Shaffer, C.D.; MacIntyre, R.J.; Roush, R.T.
Proc. Natl. Acad. Sci. U.S.A. 88; 7209-7213, 1991
A.Molecule: Molecular cloning and transformation of cyclodiene resistance in Drosophila: an
A.Reference number: A11145; MUID:91334435
A.Accession: A11145
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A.Molecule type: mRNA
A.Request: 1-606 <FFR>
A.Cross-references: GB:M69057
C:Genetics:
A:Gene: FlyBase:Rd1
A:Cross-references: FlyBase:FBgn0004244
A:Superfamily: G-protein coupled receptor
C:Keywords: neurotransmitter receptor; transmembrane protein

Query Match 51.0%; Score 50; DB 2; Length 606;
Best Local Similarity 77.8%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHGKHKHKKN 9
DB 42 HGHGKHKHKKN 434

RESULT 15

T13893
gene hind sight protein - fruit fly (Drosophila melanogaster)
C.Species: Drosophila melanogaster
C.Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C.Accession: T13893
R.K.M. Lankar, M.L.; Lipshitz, H.D.
Development 124; 2129-2141 1997.

A>Title: Control of germ-band retraction in Drosophila by the zinc-finger protein HIN
A:Reference number: Z17807; MUID:97330681
A:Accession: T13893
A>Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-1920 <YIP>
A:Cross-references: EMBL:U86010; NID:g2769709; PID:g2769710; PIDN:AA895640.1
C:Genetics:
A:Gene: hind sight
A:Cross-references: FlyBase:FBgn0003053
C:Function:
A:Description: probably function as a transcription factor
C:Keywords: nucleus; zinc finger

Query Match 51.0%; Score 50; DB 2; Length 1920;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HGHGKHKHKKN 9
DB 1136 HGHGKHKHKKN 1143

Search completed: July 6, 2001, 09:18:02
Job time: 648 sec

F:399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status ex
F:498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 94.9%; Score 93; DB 1; Length 621;
Best Local Similarity 93.8%; Pred. No. 8.9e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KHGHGCKHKNKGKKN 16
|||||
DB 476 KHGHGCKHKNKGKKN 491.

RESULT 4

C25486
K:kininogen, HMW precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 30-Sep-1993
C:Accession: C25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family.
A:Reference number: A92625; MUID:87137443
A:Accession: C25486
A:Molecule type: DNA
A:Residues: 1-264 <KIT>
C:Comment: The nucleotide sequence was obtained from GenBank, release 55.0.
C:Superfamily: kininogen; cystatin homology

Query Match 65.3%; Score 64; DB 2; Length 264;
Best Local Similarity 70.6%; Pred. No. 0.051;
Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

OY 2 HGHGCG--KHKNKGKKN 16
|||||
DB 120 HGHGCGRDKHTNKDKNN 136

RESULT 5

A25486
kininogen, HMW I precursor - rat
N:Contains: bradykinin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C:Accession: A25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family.
A:Reference number: A92625; MUID:87137443
A:Accession: A25486
A:Molecule type: mRNA
A:Residues: 1-639 <KIT>
A>Note: The authors translated the codon CAA for residue 347 as Asn
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-639/Product: kininogen, HMW I #status predicted <MAT>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>

Query Match 65.3%; Score 64; DB 2; Length 639;
Best Local Similarity 70.6%; Pred. No. 0.11;
Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

OY 2 HGHGCG--KHKNKGKKN 16
|||||
DB 495 HGHGCGRDKHTNKDKNN 511

RESULT 6

C27115
K:kininogen, LMW precursor - rat (fragments)

C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
C:Accession: C27115; A25488
R:Fung, W.P.; Schreiber, G.
J. Biol. Chem. 262, 9298-9308, 1987

A:Title: Structure and expression of the genes for major acute phase alpha-1-protein
A:Reference number: A92653; MUID:87250580

A:Accession: C27115
A:Molecule type: DNA

A:Residues: 1-290 <FUN>
R:Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
J. Biol. Chem. 262, 2345-2351, 1987

A:Title: Differing utilization of homologous transcription initiation sites of rat K
A:Reference number: A25488; MUID:87137465

A:Accession: A25488
A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-48 <KAG>
A:Cross-references: GB:J02662; NID:9205071; PIDN:AAA1483.1; PID:9205072
C:Superfamily: kininogen; cystatin homology

F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 61.2%; Score 60; DB 2; Length 290;
Best Local Similarity 66.7%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 HGHGCKHKNKGKKN 16
|||||
DB 173 HDHGCKHKNKGKKN 187

RESULT 7

A27115
major acute phase alpha-1 protein 1 - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A27115
R:Fung, W.P.; Schreiber, G.
J. Biol. Chem. 262, 9298-9308, 1987
A:Title: Structure and expression of the genes for major acute phase alpha-1-protein
A:Reference number: A92653; MUID:87250580
A:Accession: A27115
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-315 <FUN>
A:Genetics:
A:Gene: MAP1
C:Superfamily: kininogen; cystatin homology
F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 61.2%; Score 60; DB 2; Length 315;
Best Local Similarity 66.7%; Pred. No. 0.22;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 HGHGCKHKNKGKKN 16
|||||
DB 197 HDHGCKHKNKGKKN 211

RESULT 8

T49356
hypothetical protein BID1.100 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49356
R:Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Farthmann, B.; Holland, R.; Nyakatu
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49356

J. Biochem. 79, 1201-1222, 1976
A:Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino
A:Reference number: A91941, MUID:76260135
A:Accession: A91941
A:Molecule type: protein
A:Residues: 387-455 <HAN>
A:Notes: 398-pro, 401-Han, and 455-Iys were also found
R:Han, Y.N., Komiya, M., Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A:Reference number: A91938, MUID:75170265
A:Accession: A91938
A:Molecule type: protein
A:Residues: 456-496 <HA2>
R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A:Reference number: A92627, MUID:87137530
A:Accession: B29559
A:Molecule type: protein
A:Residues: 'Z', 20-104, 'E', 106-256, 'XX', 257-376 <SUE>
R:Loitsch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A:Reference number: A91153, MUID:86030270
A:Comments: annotation: bovine cleavage sites; bovine carbohydrate binding sites
R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A:Title: Disulfide bonds in bovine HMW kininogens.
A:Reference number: A94300
A:Contents: annotation: disulfide bonds
A:Note: article in Japanese
C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C:Comment: kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xpolptide residue is present in the kininogen prior to the release of bradykinin.
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-619/Product: HMW kininogen II #status predicted <AMT>
F:19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F:19-130/Domain: cystatin homology <CI1>
F:141-252/Domain: cystatin homology <CY2>
F:261-372/Domain: cystatin homology <CY3>
F:377-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F:378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
F:387-619/Product: HMW kininogen II light chain #status experimental <LCI>
F:418-488/Region: glycine/histidine/lysine-rich
F:119/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F:27-589, 82-93, 106-125, 141-144, 205-217, 228-247, 261-264, 325-337, 348-367/Disulfide bonds:
F:447/Binding site: carbohydrate (Asn) (covalent) #status absent
F:87, 168, 169, 204, 280/Binding site: carbohydrate (Asp) (covalent) #status experimental
F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F:197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:376-377/Cleavage site: Met-Lys (Kallikrein) #status experimental
F:380/Modified site: 4-hydroxyproline (Pro) #status predicted
F:386-387/Cleavage site: Arg-Ser (Kallikrein) #status experimental
F:396, 404, 510, 510/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:397, 398, 518, 522, 534, 546, 551, 568/Binding site: carbohydrate (Thr) (covalent) #status ex
F:496-497/Cleavage site: Arg-Val (Kallikrein) #status experimental

Query Match	94.9%	Score 93	DB 1	Length 619
Best Local Similarity	93.8%	Pred. No.	8.9e-06	
Matches 15	Conservative	Mismatches 1	Indels 0	Gaps 0
Oy	1 KAGHGCHKHKKKKNN 16			
Db	474 KKHGCHKHKKKKNN 489			
RESULT	3			

KGBOH1
kininogen, HMW I precursor - bovine
N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 22-Jun-1999
C:Accession: A01281; A91923; A91938; A29559
R:Kitamura, N.; Takekaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakinishi, S.
Nature 305, 543-549, 1983
A:Title: A single gene for bovine high molecular weight and low molecular weight kinin
A:Reference number: A93317; MUID:84014106
A:Accession: A01281
A:Molecule type: mRNA
A:Residues: 1-621 <KTI>
A:Cross-references: GB:V01491; GB:K01757; NID:9491; PIDN:CAA24735.1; PID:9492
R:Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
A:Reference number: A91923; MUID:70180420
A:Accession: A91923
A:Molecule type: protein
A:Residues: 378-393 <KAT>
R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen.
A:Reference number: A91938; MUID:75170265
A:Accession: A91938
A:Molecule type: protein
A:Residues: 458-498 <KAN>
R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga
J. Biol. Chem. 262, 2768-2779, 1987
A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions o
A:Reference number: A92627; MUID:87137530
A:Accession: A29559
A:Molecule type: protein
A:Residues: 72, 120-123, 125-127, 129-137 <SUC>
R:Loetsch, F.; Kellermann, J.; Henschel, A.; Foerster, B.; Müller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kinin
A:Reference number: A91153; MUID:86030270
A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 36, 808, 1984
A:Title: Disulfide bonds in bovine HMW kininogens.
A:Reference number: A94300
A:Contents: annotation; disulfide bonds
A>Note: article in Japanese
C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is i
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
X:Protein residue is present in the kininogen prior to the release of bradykinin.
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-621/Product: HMW prokininogen I #status predicted <KAT>
F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F:19-130/Domain: cystatin homology <CY1>
F:141-252/Domain: cystatin homology <CY2>
F:263-374/Domain: cystatin homology <CY3>
F:379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F:389-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F:417-488/Region: glycine/histidine/lysine-rich
F:19/Modified site: proline carboxylic acid (Gln) (in mature form) #status experie
F:27-591,82-93,106-123,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bond
F:87,168,169,204/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F:19/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F:382/Modified site: 4-hydroxyproline (Pro) #status predicted
F:388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:398,406,512/Binding site: Arg-Asp (carboxylate) (Ser) (covalent) #status experimental

F-417-468/Region: glycine/histidine/lysine-rich #status experimental
F-119/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
F-27-591,82-93,106-125,141-144,205-211,228-247,263-266,327-339,350-369/Disulfide bonds
F-87,168,169,204/Binding site: carboxylate (Asn) (covalent) #status experimental
F-136/Binding site: carboxylate (Thr) (covalent) #status experimental
F-197/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental
F-378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F-382/Modified site: 4-Hydroxyproline (Pro) #status predicted
F-388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F-398,406,512/Binding site: carboxylate (Ser) (covalent) #status experimental

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEL2>
 R:Minidolu, T.; Carretero, O.A.; Proud, D.; Malz, D.; Schell, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88209021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A31905; MUID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAE>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plasma
 A:Reference number: A34030; MUID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
 A:Reference number: S02482; MUID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19,189-192,310-314,381-389 <LENI>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55239; MUID:95251593
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Struczek, J.; Maechli, F.; Le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabec, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; MUID:96033574
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takegaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evolu
 A:Reference number: A92345; MUID:852334583
 A:Contents: annotation; gene organization

R: Pierce, J.V.
 Fed. Proc. 27, 53-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:90255622
 A:Contents: annotation; bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is i
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 C:Comment: The residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT2>
 F:19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F:19-131/Domain: cystatin homology <CY1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:381-389/Product: bradykinin (kallidin I) #status experimental <KBDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experi
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Dissulfide bond
 F:48/Binding site: carboxylate (Asn) (covalent) #status absent
 F:169,205,294/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401,533,542,546,557,571,593,628/Binding site: carboxylate (Thr) (covalent) #status
 F:577/Binding site: carboxylate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 98; DB 1; Length 644;
 Best local similarity 100.0%; Pred. No. 1.8e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHGHGKHKHKKGN 16
 DB 498 KHGHGKHKHKKGN 513
 |||||

RESULT 2
 KGBH2
 kininogen, HMW II precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 M:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence revision 14-Nov-1993 #text change 22-Jun-1999
 C:Accession: A01282; A91923; A91941; A91938; B29559
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A:Title: A single gene for bovine high molecular weight and low molecular weight kin
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01282
 A:Molecule type: mRNA
 A:Residues: 1-619 <KIT>
 A:Cross-references: GB:V01492; GB:K01758; NID:9493; PIDN:CAA24736.1; PID:9494
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 376-391 <KAT>
 R:Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:18:02 ; Search time 73.59 Seconds
(without alignments)
16.562 Million cell updates/sec

Title: US-09-437-912-8
Perfect score: 98
Sequence: 1 KHCHGKHKNNKGN 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	644	1 KGHUHI	kininogen, HMW pre
2	93	94.9	619	1 KGBOH1	kininogen, HMW II
3	93	94.9	621	1 KGBOH1	kininogen, HMW I P
4	64	65.3	264	2 C25486	K-kininogen, HMW P
5	64	65.3	639	2 A25486	kininogen, HMW I P
6	60	61.2	290	2 C27115	K-kininogen, LMW P
7	60	61.2	315	2 A27115	major acute phase
8	54.5	55.6	125	2 T49356	hypothetical prote
9	54.5	55.1	1891	2 T13594	hypothetical prote
10	51	52.0	359	2 S14283	transcription fact
11	50	51.0	98	2 S08137	gene 2C protein -
12	50	51.0	229	2 T27840	hypothetical prote
13	50	51.0	241	2 T26522	hypothetical prote
14	50	51.0	606	2 A41145	gamma-aminobutyric
15	50	51.0	1920	2 T13893	gene handstight pro
16	49.5	50.5	140	2 T27059	hypothetical prote
17	48.5	49.5	85	2 A45969	hemolymph anti-fung
18	48.5	49.5	199	2 T48099	hypothetical prote
19	48	49.0	110	2 T07618	cold stress prote
20	48	49.0	207	2 T08109	oleosin-like prote
21	48	49.0	251	2 T34168	hypothetical prote
22	48	49.0	306	2 S08607	chorion protein s3
23	48	49.0	375	2 T08134	oleosin-like prote
24	48	49.0	480	2 H64789	probable 2-compone
25	48	49.0	482	2 C85555	gene pipsqueak pro
26	48	49.0	535	2 S66148	gene pipsqueak pro
27	47.5	48.5	1085	2 S66149	zinc transporter 2
28	47.5	48.5	503	2 S54302	zinc transporter 2
29	47.5	48.5	507	2 S54303	zinc transporter pro

30	47	48.0	98	2 JC5822	NADH dehydrogenase
31	47	48.0	208	2 T08132	oleosin-like prote
32	47	48.0	1037	2 D96786	protein P10A5.15 I
33	46.5	47.4	137	2 S65780	glycine/proline-ri
34	46	46.9	136	2 JQ2266	cold acclimation p
35	46	46.9	332	2 T29574	hypothetical prote
36	46	46.9	378	2 S71224	xyloglucan endo-1,
37	46	46.9	332	2 T49164	zinc transporter-1
38	46	46.9	391	2 T26756	hypothetical prote
39	46	46.9	410	2 T26757	hypothetical prote
40	46	46.9	529	2 T08684	hypothetical prote
41	46	46.9	549	2 T15506	hypothetical prote
42	46	46.9	1002	2 T43236	carboxypeptidase C
43	46	46.9	1213	2 A54063	TARA-binding prote
44	45.5	46.4	179	2 A85217	hypothetical prote
45	45.5	46.4	439	2 S58327	cobalt accumulatio

ALIGNMENTS

RESULT 1
KGHUI
kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N:Contains: bradykinin (Kallidin I); HMW kininogen I; HMW kininogen II; low molecular
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence revision 28-May-1986 #text change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27899; A31905; A34030;
R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A:Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:85122621
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:M1437; NID:9177889
R:Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and 1
A:Reference number: A92544; MUID:85234582
A:Accession: A25276
A:Molecule type: mRNA
A:Residues: 1-592, 'I', 594-644 <TAK>
A:Cross-references: GB:M1437; NID:9186751; PIDN:AAB59550.1; PID:9386852
R:Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A:Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:9323854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 1-NSM', 253-377 <AUE>
A:Note: differences are due to known cloning artifacts
R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foerlsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kin
A:Reference number: A91153; MUID:86030270
A:Accession: A91153
A:Molecule type: protein
A:Residues: 379-644 <LOT>
A:Note: the bradykinin sequence preceding the light chain sequence was not determined
R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Muller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A:Title: Completion of the primary structure of human high-molecular-mass kininogen.
A:Reference number: A24871; MUID:86108361
A:Accession: A24871
A:Molecule type: protein
A:Residues: 'Z', 20-380 <KEL>
R:Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
In: Kinins IV, Greenbaum, L.M., and Margolis, H.S., ed., pp.85-89, Plenum Press, New
York, 1985
A:Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A:Reference number: A27899
A:Accession: A27899

Db 355 KGGGKGGKGGK 368

RESULT 13

US-08-933-774-7
Sequence 7, Application US/08933774A
Patent No. 6025137

GENERAL INFORMATION:

APPLICANT: Shyjan, Andrew W.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS, PREVENTION
AND TREATMENT OF TUMOR PROGRESSION

FILE REFERENCE: 07334/004003

CURRENT APPLICATION NUMBER: US/08/933,774A

EARLIER FILING DATE: 1997-09-19

EARLIER APPLICATION NUMBER: US 08/623,679

EARLIER FILING DATE: 1996-03-29

EARLIER APPLICATION NUMBER: US 08/412,431

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 7

LENGTH: 1497

TYPE: PRT

ORGANISM: Homo sapiens

US-08-933-774-7

Query Match

Best Local Similarity 44.9%; Score 44; DB 3; Length 1497;
Pred. No. 1.6e+02; Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KHGHGKHKHKKG 14

Db 355 KGGGKGGKGGK 368

RESULT 14

US-08-623-679-9
Sequence 9, Application US/08623679
Patent No. 5674739

GENERAL INFORMATION:

APPLICANT: Shyjan, Andrew W.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
DIAGNOSIS, PREVENTION AND TREATMENT OF TUMOR

TITLE OF INVENTION: PROGRESSION

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/623,679

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/412,431

FILING DATE: 29-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Fasse, J. Peter

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 07334/004001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1533 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-623-679-9

Query Match

Best Local Similarity 44.9%; Score 44; DB 1; Length 1533;
Pred. No. 1.6e+02; Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KHGHGKHKHKKG 14

Db 355 KGGGKGGKGGK 368

RESULT 15

US-08-933-774-9
Sequence 9, Application US/08933774A
Patent No. 6025137

GENERAL INFORMATION:

APPLICANT: Shyjan, Andrew W.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS, PREVENTION
AND TREATMENT OF TUMOR PROGRESSION

FILE REFERENCE: 07334/004003

CURRENT APPLICATION NUMBER: US/08/933,774A

EARLIER FILING DATE: 1997-09-19

EARLIER APPLICATION NUMBER: US 08/623,679

EARLIER FILING DATE: 1996-03-29

EARLIER APPLICATION NUMBER: US 08/412,431

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 9

LENGTH: 1533

TYPE: PRT

ORGANISM: Homo sapiens

US-08-933-774-9

Query Match

Best Local Similarity 44.9%; Score 44; DB 3; Length 1533;
Pred. No. 1.6e+02; Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KHGHGKHKHKKG 14

Db 355 KGGGKGGKGGK 368

Search completed: July 6, 2001, 09:10:24
Job time: 190 sec

ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,080
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-28

Query Match 44.9%; Score 44; DB 1; Length 1085;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KHGHGKHKNG 13
| | | :||||
Db 521 KSGKSHHKNG 533

RESULT 11
US-08-938-534-28
Sequence 28, Application US/08938534
Patent No. 5916752
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,534
FILING DATE: 26-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,080
FILING DATE:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-534-28

Query Match 44.9%; Score 44; DB 2; Length 1085;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KHGHGKHKNG 13
| | | :||||
Db 521 KSGKSHHKNG 533

RESULT 12
US-08-623-679-7
Sequence 7, Application US/08623679
Patent No. 5674739
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: DIAGNOSIS, PREVENTION AND TREATMENT OF TUMOR
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/623,679
FILING DATE: 29-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/412,431
FILING DATE: 29-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07334/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1497 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-623-679-7

Query Match 44.9%; Score 44; DB 1; Length 1497;
Best Local Similarity 64.3%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KHGHGKHKNG 14

US-09-208-742-2

Query Match 46.9%; Score 46; DB 4; Length 1199;

Best Local Similarity 53.3%; Pred. No. 70;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 HGHGSHKHNKGN 16

Db 1147 HHHHHHKKKKKK 1161

RESULT 8

US-08-188-582-20

; Sequence 20, Application US/08188582

; Patent No. 5534410

; GENERAL INFORMATION:

; APPLICANT: Tjian, Robert

; APPLICANT: Comai, Lucio

; APPLICANT: Dynlacht, Brian D.

; APPLICANT: Hoey, Timothy

; APPLICANT: Ruppert, Siegfried

; APPLICANT: Tanese, Naoko

; APPLICANT: Wang, Edith

; APPLICANT: Weinzierl, Robert O.J.

; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,

; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/188, 582

; FILING DATE: 28-JAN-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman, Richard A.

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1213 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-188-582-20

Query Match 46.9%; Score 46; DB 1; Length 1213;

Best Local Similarity 50.0%; Pred. No. 70;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KHGHGSHKHNKGN 16

Db 1163 KKHHRSHKDKERKD 1178

RESULT 9

US-08-646-715-20

; Sequence 20, Application US/08646715

; Patent No. 5637686

; GENERAL INFORMATION:

; APPLICANT: Tjian, Robert

; APPLICANT: Comai, Lucio

; APPLICANT: Dynlacht, Brian D.

; APPLICANT: Hoey, Timothy

; APPLICANT: Ruppert, Siegfried

; APPLICANT: Tanese, Naoko

; APPLICANT: Wang, Edith

; APPLICANT: Weinzierl, Robert O.J.

; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,

; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/646, 715

; FILING DATE: 09-MAY-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/188, 582

; FILING DATE: 28-JAN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman, Richard A.

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1213 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-646-715-20

Query Match 46.9%; Score 46; DB 1; Length 1213;

Best Local Similarity 50.0%; Pred. No. 70;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KHGHGSHKHNKGN 16

Db 1163 KKHHRSHKDKERKD 1178

RESULT 10

US-08-431-080-28

; Sequence 28, Application US/08431080

; Patent No. 5698686

; GENERAL INFORMATION:

; APPLICANT: Gottschling, Daniel E.

; APPLICANT: Singer, Miriam S.

; TITLE OF INVENTION: Telomerase Compositions and Methods

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P. O. Box 4433

; CITY: Houston

; STATE: TEXAS

; COUNTRY: UNITED STATES OF AMERICA

TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-072-064-6

Query Match 51.0%; Score 50; DB 3; Length 637;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHGHGKH 9
:|||||
DB 457 EHGHGHH 465

RESULT 5

US-08-072-064-8
Sequence 8, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FERRENCH-CONSTANT, RICHARD H.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,064
FILING DATE: 19930602
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 770,881
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OPND-00574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-072-064-8

Query Match 51.0%; Score 50; DB 3; Length 637;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHGHGKH 9
:|||||
DB 457 EHGHGHH 465

RESULT 6

PCT-US92-08558-1
Sequence 1, Application PC/TUS9208558
GENERAL INFORMATION:
APPLICANT: Corneil Research Foundation, Inc.
TITLE OF INVENTION: MOLECULAR CLONING AND TRANSFORMATION OF CYCLODIENE RESISTAN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08558
FILING DATE: 19921002
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/770,881
FILING DATE: October 4th 1991
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: CRF D-1052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster
POSITION IN GENOME:
CHROMOSOME/SEGMENT: III
MAP POSITION: approximately map unit 26
PCT-US92-08558-1

Query Match 51.0%; Score 50; DB 5; Length 637;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHGHGKH 9
:|||||
DB 457 EHGHGHH 465

RESULT 7

US-09-208-742-2
Sequence 2, Application US/09208742
Patent No. 6174679
GENERAL INFORMATION:
APPLICANT: Kaufmann, Joerg
TITLE OF INVENTION: CIP150/HTAF1150 is Necessary for Cell
FILE REFERENCE: 1453.002
CURRENT APPLICATION NUMBER: US/09/208,742
CURRENT FILING DATE: 1998-12-10
NUMBER OF SEQ. ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 2
LENGTH: 1199
TYPE: PRT
ORGANISM: human

```
RESULT 2
US-08-072-064-1
; Sequence 1, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Drosophila melanogaster
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: III; polytene subregion 66F
; MAP POSITION: approximately map unit 26
US-08-072-064-1

Query Match          51.0%; Score 50: DB 3; Length 637;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHGHGHGH 9
:|||||
Db 457 EHGHGHGH 465

RESULT 3
US-08-072-064-4
; Sequence 4, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
```

```
STATE: California
COUNTRY: United States of America
ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-072-064-4
```

```
Query Match          51.0%; Score 50: DB 3; Length 637;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHGHGHGH 9
:|||||
Db 457 EHGHGHGH 465
```

```
RESULT 4
US-08-072-064-6
; Sequence 6, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPD-00574
; TELECOMMUNICATION INFORMATION:
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:23 ; Search time 56.74 Seconds
(without alignments)
5.681 Million cell updates/sec

Title: US-09-437-912-8

Perfect score: 98
Sequence: 1 KHGHGKHKNNKKKN 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	51.0	617	1 US-08-137-614A-26	Sequence 26, App1
2	50	51.0	637	3 US-08-072-064-1	Sequence 1, App1
3	50	51.0	637	3 US-08-072-064-4	Sequence 4, App1
4	50	51.0	637	3 US-08-072-064-6	Sequence 6, App1
5	50	51.0	637	3 US-08-072-064-8	Sequence 8, App1
6	50	51.0	637	5 PCT-US92-08558-1	Sequence 1, App1
7	46	46.9	1199	4 US-09-208-742-2	Sequence 2, App1
8	46	46.9	1213	1 US-08-188-582-20	Sequence 20, App1
9	46	46.9	1213	1 US-08-646-715-20	Sequence 20, App1
10	44	44.9	1085	1 US-08-431-080-28	Sequence 28, App1
11	44	44.9	1085	2 US-08-938-534-28	Sequence 28, App1
12	44	44.9	1497	1 US-08-623-679-7	Sequence 7, App1
13	44	44.9	1497	3 US-08-933-774-7	Sequence 7, App1
14	44	44.9	1533	1 US-08-623-679-9	Sequence 9, App1
15	44	44.9	1533	3 US-08-933-774-9	Sequence 9, App1
16	43	43.9	515	4 US-08-942-0128-32	Sequence 32, App1
17	43	43.9	599	1 US-08-172-331B-4	Sequence 4, App1
18	43	43.9	599	2 US-09-032-315-6	Sequence 6, App1
19	43	43.9	599	2 US-08-993-318A-6	Sequence 6, App1
20	43	43.9	599	4 US-09-399-886-6	Sequence 6, App1
21	43	43.9	1382	4 US-09-396-260-6	Sequence 6, App1
22	43	43.9	1657	3 US-09-057-570-4	Sequence 4, App1
23	43	43.9	1657	3 US-09-057-570-2	Sequence 2, App1
24	43	43.9	1805	3 US-09-057-570-7	Sequence 7, App1
25	42	42.9	339	2 US-08-758-621-2	Sequence 2, App1
26	42	42.9	339	4 US-09-107-858-2	Sequence 2, App1
27	42	42.9	456	2 US-08-709-979A-1	Sequence 1, App1

28	42	42.9	456	4 US-08-709-979A-11	Sequence 11, App1
29	41	41.8	388	2 US-08-382-505-2	Sequence 2, App1
30	41	41.8	542	1 US-08-412-431-3	Sequence 3, App1
31	41	41.8	542	1 US-08-623-679-3	Sequence 3, App1
32	41	41.8	542	3 US-08-933-774-3	Sequence 3, App1
33	41	41.8	903	1 US-08-750-532-1	Sequence 1, App1
34	41	41.8	1398	1 US-08-750-532-9	Sequence 9, App1
35	40.5	41.3	800	1 US-08-785-052-4	Sequence 4, App1
36	40.5	41.3	800	2 US-08-913-581-4	Sequence 4, App1
37	40	40.8	13	1 US-07-840-077A-6	Sequence 6, App1
38	40	40.8	13	1 US-08-484-184-2	Sequence 2, App1
39	40	40.8	13	1 US-08-454-950-6	Sequence 6, App1
40	40	40.8	13	1 US-08-087-219-2	Sequence 2, App1
41	40	40.8	13	1 US-08-269-929-6	Sequence 6, App1
42	40	40.8	13	1 US-08-454-949-6	Sequence 6, App1
43	40	40.8	14	1 US-07-694-983-11	Sequence 11, App1
44	40	40.8	15	1 US-07-694-983-19	Sequence 19, App1
45	40	40.8	20	1 US-07-694-983-12	Sequence 12, App1

ALIGNMENTS

RESULT 1
US-08-137-614A-26
; Sequence 26, Application US/08137614A
; Patent No. 5487976
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipfle, Douglas C.
; APPLICANT: Henderson, Joseph E.
; TITLE OF INVENTION: Gene Encoding An Insect
; TITLE OF INVENTION: Gamma-Aminobutyric Acid (GABA) Receptor Subunit
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,614A
; FILING DATE: 15-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)263-1636
; TELEFAX: (716)263-1600
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 617 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-137-614A-26

Query Match 51.0%; Score 50; DB 1; Length 617;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 KHGHGKHK 9
DB 436 EHGHGKHH 444

DE CG11718 PROTEIN.
GN CG11718.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SROUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Modarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003678; AAF54238.1; -;
DR FLYBase: FBgn0037585; CG11718.
SQ SEQUENCE 819 AA; 92141 MW; E7F295E74FE2A72B CRC64;

Query Match 52.0%; Score 51; DB 5; Length 819;
Best Local Similarity 55.6%; Pred. No. 6.4;
Matches 10; Conservative 2; Mismatches 4; Indels 2; Gaps 1;
QY 1 KHGHH--CKHKKGRKN 16
DB 292 KHGKHKKHKKHKKSSS 309

Search completed: July 6, 2001, 09:25:56
Job time: 992 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:24 ; Search time 56.74 Seconds
(without alignments)
5.681 Million cell updates/sec

Title: US-09-437-912-9

Perfect score: 98
Sequence: 1 HKNKGKNGKNGKNGKT 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents-AA:
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42.5	43.4	1507	6 5268270-2	Patent No. 5268270
2	41	41.8	793	2 US-08-468-558-5	Sequence 5, Appl
3	41	41.8	1382	2 US-08-737-715-2	Sequence 2, Appl
4	40	40.8	376	1 US-08-594-031-100	Sequence 100, App
5	40	40.8	376	1 US-08-594-031-102	Sequence 102, App
6	40	40.8	422	3 US-08-872-979-8	Sequence 8, Appl
7	39	39.8	60	2 US-08-117-952-787	Sequence 787, App
8	39	39.8	60	2 US-08-117-952-788	Sequence 788, App
9	39	39.8	3177	2 US-08-477-451-4	Sequence 4, Appl
10	39	39.8	3898	2 US-08-750-717-2	Sequence 2, Appl
11	38.5	39.3	606	2 US-08-883-534-3	Sequence 3, Appl
12	38.5	39.3	172	4 US-09-204-764-3	Sequence 3, Appl
13	38	38.8	172	4 US-08-916-576B-4	Sequence 4, Appl
14	38	38.8	293	3 US-09-203-716-2	Sequence 12, Appl
15	38	38.8	345	1 US-08-183-214-12	Sequence 30, Appl
16	37	37.8	78	2 US-07-885-089B-30	Sequence 33, Appl
17	37	37.8	78	2 US-07-885-089B-33	Sequence 33, Appl
18	37	37.8	83	2 US-07-885-089B-16	Sequence 18, Appl
19	37	37.8	84	6 5202428-10	Patent No. 5202428
20	37	37.8	84	6 5202428-10	Sequence 1, Appl
21	37	37.8	91	4 US-09-077-977A-1	Sequence 40, Appl
22	37	37.8	95	2 US-08-484-438-40	Patent No. 5514582
23	37	37.8	126	6 5514582-43	Patent No. 5514582
24	37	37.8	137	6 5202428-9	Sequence 37, Appl
25	37	37.8	146	2 US-07-885-089B-37	Sequence 36, Appl
26	37	37.8	152	2 US-07-885-089B-36	Sequence 34, Appl
27	37	37.8	158	2 US-07-885-089B-34	Sequence 34, Appl

28	37	37.8	192	1 US-08-208-008C-9	Sequence 9, Appl
29	37	37.8	226	2 US-07-885-089B-35	Sequence 35, Appl
30	37	37.8	247	2 US-07-885-089B-2	Sequence 8, Appl
31	37	37.8	247	2 US-07-885-089B-8	Sequence 2, Appl
32	37	37.8	252	2 US-07-885-089B-7	Sequence 2, Appl
33	37	37.8	263	1 US-07-906-983-2	Sequence 1, Appl
34	37	37.8	286	3 US-09-203-716-1	Sequence 2, Appl
35	37	37.8	352	1 US-08-785-052-2	Sequence 2, Appl
36	37	37.8	352	2 US-08-913-581-2	Sequence 2, Appl
37	37	37.8	384	1 US-07-783-706-2	Sequence 2, Appl
38	37	37.8	384	2 US-08-445-342A-2	Sequence 2, Appl
39	37	37.8	384	5 US-09-066-481-2	Sequence 2, Appl
40	37	37.8	384	5 PCT-US92-09124-2	Sequence 2, Appl
41	37	37.8	462	2 US-08-484-438-42	Sequence 42, Appl
42	37	37.8	533	2 US-08-225-488-2	Sequence 3, Appl
43	37	37.8	542	1 US-08-412-431-3	Sequence 3, Appl
44	37	37.8	542	1 US-08-623-679-3	Sequence 3, Appl
45	37	37.8	542	3 US-08-933-774-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
5268270-2
Patent No. 5268270
APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohner, Johannes
TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
NEGATIVE HOST CELLS
NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/171,872
FILING DATE: 01-JUL-1987
SEQ ID NO: 2:
LENGTH: 1507
5268270-2

Query Match 43.4%; Score 42.5; DB 6; Length 1507;
Best Local Similarity 55.6%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 2;

QY 1 HKNKG--KKNKGKNGKKT 16
DB 301 HDNAGTVKNGEHH-WKT 317

RESULT 2
US-08-468-558-5
Sequence 5, Application US/08468558
Patent No. 5877280
GENERAL INFORMATION:
APPLICANT: Metmur, James G.
TITLE OF INVENTION: Cloning and Expression of Thermostable
TITLE OF INVENTION: Muts Genes and Proteins and Uses Therefor
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
CITY: Lexington
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,558
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MSM94-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-9540
TELEFAX: 617-861-6240
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-558-5

Query Match 41.8%; Score 41; DB 2; Length 793;
Best Local Similarity 72.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KNGKNGKNGK 12
Db 772 KNTKNGKNGKSN 782

RESULT 3
US-08-737-715-2
Sequence 2, Application US/08737715
Patent No. 5958685
GENERAL INFORMATION:
APPLICANT: Edina, Yousuke
TITLE OF INVENTION: MUTANT HUMAN INSULIN RECEPTOR DNA
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MIOT, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,715
FILING DATE: 12-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 134827/1995
FILING DATE: 12-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Dean H.
REGISTRATION NUMBER: 33,981
REFERENCE/DOCKET NUMBER: Q-43323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1382 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-737-715-2

Query Match 41.8%; Score 41; DB 2; Length 1382;
Best Local Similarity 70.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 HKNGKNGK 10
I I I I I I I I

Db 1363 HNGKNGKNGR 1372

RESULT 4
US-08-594-031-100
Sequence 100, Application US/08594031
Patent No. 5763182
GENERAL INFORMATION:
APPLICANT: THOMPSON, Timothy C.
TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 1299 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004-2400
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,031
FILING DATE: 30-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-594-031-100

Query Match 40.8%; Score 40; DB 1; Length 376;
Best Local Similarity 62.5%; Pred. No. 92;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 NGKNGK 15
Db 114 DGRHNGWR 121

RESULT 5
US-08-594-031-102
Sequence 102, Application US/08594031
Patent No. 5763182
GENERAL INFORMATION:
APPLICANT: THOMPSON, Timothy C.
TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 1299 Pennsylvania Avenue, N.W.
CITY: Washington

STATE: DC
COUNTRY: USA
ZIP: 20004-2400
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,031
FILING DATE: 30-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-594-031-102

Query Match 40.8%; Score 40; DB 1; Length 376;
Best local Similarity 62.5%; Pred. No. 92;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 NGKHNGWK 15
DB 114 DSKHNGWR 121

RESULT 6
US-08-872-979-8
Sequence 8, Application US/08872979
Patent No. 6074844
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
SOFTWARE: Corley, Neil C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,979
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0320 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 338658
US-08-872-979-8

Query Match 40.8%; Score 40; DB 3; Length 422;
Best local Similarity 57.1%; Pred. No. 16+02;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 KNGKNGKHNGWK 15
DB 86 KKKKKKKKKGCKG 99

RESULT 7
US-08-117-952-787
Sequence 787, Application US/08117952
Patent No. 5851760
GENERAL INFORMATION:
APPLICANT: Evans, Glen A.
APPLICANT: Smith, Michael W.
TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
TITLE OF INVENTION: SAMPLED MAPS OF COMPLEX GENOMES
NUMBER OF SEQUENCES: 797
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,952
FILING DATE: 07-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/078,471
FILING DATE: 15-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9423
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 787:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein

FRAGMENT TYPE: Internal
US-08-117-952-787

Query Match 39.8%; Score 39; DB 2; Length 60;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

QY 5 GKKGKNGHNG---WK 15
||||| | | | |
DB 10 GKKNSPHEGKRIMWK 24

RESULT 8
US-08-117-952-788
Sequence 788, Application US/08117952
Patent No. 5851760

GENERAL INFORMATION:
APPLICANT: Evans, Glen A.
APPLICANT: Smith, Michael W.
TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
TITLE OF INVENTION: SAMPLED MAPS OF COMPLEX GENOMES
NUMBER OF SEQUENCES: 797
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Bueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,952
CLASSIFICATION: 435
FILING DATE: 07-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/078,471
FILING DATE: 15-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9423
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 788:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-117-952-788

Query Match 39.8%; Score 39; DB 2; Length 60;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

QY 5 GKKGKNGHNG---WK 15
||||| | | | |
DB 10 GKKNSPHEGKRIMWK 24

RESULT 9
US-08-477-451-4
Sequence 4, Application US/08477451
Patent No. 5928865
GENERAL INFORMATION:

APPLICANT: Covacci, Antonello
TITLE OF INVENTION: Helicobacter Pylori CagI Region
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,451
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0335.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2708
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3177 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-451-4

Query Match 39.8%; Score 39; DB 2; Length 3177;
Best Local Similarity 60.0%; Pred. No. 1,1e+03;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 KKGKNGKWK 15
||||| | | | |
DB 910 KKNRYNGYK 919

RESULT 10
US-08-750-717-2
Sequence 2, Application US/08750717
Patent No. 6180109
GENERAL INFORMATION:
APPLICANT: MOORMANN, Robertus J. M.
APPLICANT: VAN RIJN, Petrus A.
TITLE OF INVENTION: Nucleotide Sequences of Pestivirus
TITLE OF INVENTION: Strains, Polypeptides Encoded by These Sequences and Use
TITLE OF INVENTION: Theeef for Diagnosis and Prevention of Pestivirus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,717
FILING DATE: 24-DEC-1996
CLASSIFICATION: 424

```

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: EP 94201743.5
3 FILING DATE: 17-JUN-1994
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: WO PCT/ML95/00214
6 FILING DATE: 16-JUN-1995
7 ATTORNEY/AGENT INFORMATION:
8 NAME: PATCH, Andrew J.
9 REGISTRATION NUMBER: 32,925
10 REFERENCE/DOCKET NUMBER: BO 39123
11 TELECOMMUNICATION INFORMATION:
12 TELEPHONE: 703-521-2297
13 TELEFAX: 703-685-0573
14
15 INFORMATION FOR SEQ ID NO: 2:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 3898 amino acids
18 TYPE: amino acid
19 TOPOLOGY: linear
20 MOLECULE TYPE: protein
21 US-08-750-717-2
22
23 Query Match 39.8%; Score 39; DB 4; Length 3898;
24 Best Local Similarity 44.4%; Pred. No. 1.3e+03;
25 Matches 8; Conservative 1; Mismatches 5; Indels 4; Gaps 1
26
27 QY 1 HKNK---GKKNGKHNGW 14
28 ||| | | | | |
29 Db 3262 HKMKTLCPGSSLGRRNDW 3279
30
31 RESULT 11
32 US-08-883-534-3
33 Sequence 3, Application US/08883534
34 Patent No. 5846777
35 GENERAL INFORMATION:
36 APPLICANT: Bandman, Olga
37 APPLICANT: Lal, Preeti
38 APPLICANT: Corley, Neil C.
39 TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
40 NUMBER OF SEQUENCES: 6
41 CORRESPONDENCE ADDRESS:
42 ADDRESSEE: Incyte Pharmaceuticals, Inc.
43 STREET: 3174 Porter Drive
44 CITY: Palo Alto
45 STATE: CA
46 COUNTRY: USA
47 ZIP: 94304
48 COMPUTER READABLE FORM:
49 MEDIUM TYPE: Diskette
50 COMPUTER: IBM Compatible
51 OPERATING SYSTEM: DOS
52 SOFTWARE: FastSeq for Windows Version 2.0
53 CURRENT APPLICATION DATA:
54 APPLICATION NUMBER: US/08/883,534
55 FILING DATE:
56 CLASSIFICATION: 424
57 PRIOR APPLICATION DATA:
58 APPLICATION NUMBER:
59 FILING DATE:
60 ATTORNEY/AGENT INFORMATION:
61 NAME: Billings, Lucy J.
62 REGISTRATION NUMBER: 36,749
63 REFERENCE/DOCKET NUMBER: PF-0332 US
64 TELECOMMUNICATION INFORMATION:
65 TELEPHONE: 415-855-0555
66 TELEFAX: 415-845-4166
67
68 TELEX:
69 INFORMATION FOR SEQ ID NO: 3:
70 SEQUENCE CHARACTERISTICS:
71 LENGTH: 606 amino acids
72 TYPE: amino acid

```

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: NEUTGMT01
: CLONE: 1221143
: US-08-883-534-3

Query Match
Best Local Similarity 39.3%; Score 38.5; DB 2; Length 606;
Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 1

QY 1 HKNGKR--NGKNG 13
||| || :||:|
Db 329 HKNGKSYISGSHDG 344

RESULT 12
US-09-204-764-3
: Sequence 3, Application US/09204764
: Patent No. 6025464
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Lal, Preeti
: TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/204,764
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/883,534
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0332 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 606 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: NEUTGMT01
: CLONE: 1221143
: US-09-204-764-3

Query Match
Best Local Similarity 39.3%; Score 38.5; DB 3; Length 606;
Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 1

```

RESULT 13
US-08-916-576B-4
; Sequence 4, Application US/08916576B
; Patent No. 6171816
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; APPLICANT: DILLON, PATRICK J.
; APPLICANT: EBNER, REINHARD
; APPLICANT: ENDRESS, GREGORY A.
; TITLE OF INVENTION: NOVEL HUMAN GROWTH FACTORS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,576B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,347
; FILING DATE: 23-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0500001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 172 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-916-576B-4

Query Match 38.8%; Score 38; DB 4; Length 172;
Best Local Similarity 43.8%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 HNNKGGKKNKGHWKT 16
DB 27 HNGLGKFGDHIWRT 42

RESULT 14
US-09-203-716-2
; Sequence 2, Application US/09203716
; Patent No. 6001653
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
; FILE REFERENCE: ISH-0333
; CURRENT APPLICATION NUMBER: US/09/203,716
; CURRENT FILING DATE: 1998-12-02
; EARLIER APPLICATION NUMBER: 60/067,458
; EARLIER FILING DATE: 1997-12-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-203-716-2

Query Match 38.8%; Score 38; DB 3; Length 293;
Best Local Similarity 60.0%; Pred. No. 14+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 KNGKHGKWT 16
DB 230 ENKKTNGWRT 239

RESULT 15
US-08-183-214-12
; Sequence 12, Application US/08183214
; Patent No. 5716816
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Stanley, Sally J.
; APPLICANT: Nightingale, Maria S.
; APPLICANT: Murtagh, Jr., James J.
; APPLICANT: Monaco, Lucia
; APPLICANT: Takada, Tatsuyuki
; TITLE OF INVENTION: CLONES ENCODING MAMMALIAN
; TITLE OF INVENTION: ADP-RIBOSYLARGININE HYDROLASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,214
; FILING DATE: 14-JAN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,231
; FILING DATE: 22-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-183-214-12

Query Match 38.8%; Score 38; DB 1; Length 345;
Best Local Similarity 46.7%; Pred. No. 1.7+02;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 HNNKGGKKNKGHWKT 15
DB 93 HNNMQLKPGKPNWKT 107

Fri Jul 6 09:48:43 2001

us-09-437-912-9.ra1

Page 7

Search completed: July 6, 2001, 09:10:24
Job time: 190 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:18:02 ; Search time 73.59 seconds

(without alignments)
16.562 Million cell updates/sec

Title: US-09-437-912-9

Sequence: 1 HKKGGKKKNGKNGMKT 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

1: PIR68:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	644	1	KGHUH1
2	75	76.5	619	1	KGBOH2
3	75	76.5	621	1	KGBOH1
4	47	48.0	225	2	T39083
5	47	48.0	275	2	T35064
6	47	48.0	344	2	S34153
7	46	46.9	169	2	G71333
8	46	46.9	516	2	T49422
9	46	46.9	648	1	H69878
10	46	46.9	857	2	T05352
11	45	45.9	601	2	T51748
12	44	44.9	163	1	S21633
13	44	44.9	298	2	A84100
14	44	44.9	645	2	I50680
15	43	43.9	81	2	A86772
16	43	43.9	174	2	S73219
17	43	43.9	374	2	T11662
18	43	43.9	504	2	B71620
19	43	43.9	808	2	G86185
20	42.5	43.4	1532	2	A26039
21	42	42.9	83	2	A05157
22	42	42.9	678	2	A54514
23	42	42.9	746	2	T05899
24	42	42.9	799	2	T34090
25	42	42.9	1009	2	T31081
26	42	42.9	1230	2	T04181
27	41.5	42.3	316	2	T25179
28	41	41.8	145	2	F81807
29	41	41.8	145	2	H81061

30	41	41.8	186	2	G71812
31	41	41.8	303	2	T06618
32	41	41.8	311	2	E86746
33	41	41.8	359	1	H64688
34	41	41.8	359	2	F71827
35	41	41.8	375	2	T08134
36	41	41.8	385	2	S72275
37	41	41.8	552	2	D81290
38	41	41.8	609	1	G69843
39	41	41.8	628	2	T44581
40	41	41.8	645	2	T26926
41	41	41.8	671	2	JF0288
42	41	41.8	678	2	T32483
43	41	41.8	793	2	C72219
44	41	41.8	812	1	IS2PT1
45	41	41.8	814	2	T50327

ALIGNMENTS

RESULT 1

KGHUH1
kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030;
R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shikawa, H.; Sasaki, M.
Biochemistry 23: 5691-5697, 1984
A:Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:8512621
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:K02566; NID:9177889
R:Takagaki, Y.; Kitamura, N.; Nakamishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and 1
A:Reference number: A92544; MUID:85234582
A:Accession: A25276
A:Molecule type: mRNA
A:Residues: 1-592, 594-644 <TRK>
A:Cross-references: GB:M1437; NID:9186751; PIDN:AAB59550.1; PID:9386852
R:Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A:Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:93223854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 'ANSM', 253-377 <AUE>
R:Lotspeltch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kint
A:Reference number: A91153; MUID:86030270
A:Accession: A91153
A:Molecule type: protein
A:Residues: 379-644 <LOT>
A>Note: the bradykinin sequence preceding the light chain sequence was not determined
R:Kellermann, J.; Lotspeltch, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A:Title: Completion of the primary structure of human high-molecular-mass kininogen.
A:Reference number: A24871; MUID:86108361
A:Accession: A24871
A:Molecule type: protein
A:Residues: 'Z', 20-380 <KEU>
R:Kellermann, J.; Lotspeltch, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
A:Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A:Reference number: A27899
A:Accession: A27899

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEU2>
 A:Title: Structural features of plasma kinins and kininogens.
 R:Androu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; PMID:88209021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIV>
 R:Mada, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A31905; PMID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAE>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plasma
 A:Reference number: A34030; PMID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Renardic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
 A:Reference number: S02482; PMID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19, 189-192, 310-314, 381-389 <LENI>
 R:Kato, H.; Matsumura, Y.; Mada, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; PMID:8821869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Renardic, B.; Kraovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; PMID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55239; PMID:95251593
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Straczek, J.; Maachi, F.; Le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; PMID:96033974
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evolu
 A:Reference number: A95545; PMID:85234583
 A:Contents: annotation; gene organization

R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; PMID:90255622
 A:Accession: A91455
 A:Contents: annotation; bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is l
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 C:Comment: xypoline residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:RNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379, 390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F:19-131/Domain: cystatin homology <CI1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status experi
 F:28-614, 83-94, 107-126, 142-145, 206-218, 229-248, 264-267, 328-340, 351-370/Dsulfide bond
 F:48/Binding site: carboxylate (Asn) (covalent) #status absent
 F:169, 205, 294/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 593, 628/Binding site: carboxylate (Thr) (covalent) #status
 F:577/Binding site: carboxylate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 98; DB 1; Length 644;
 Best Local Similarity 100.0%; Pred. No. 8.4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKNKGNKNGHNGWKT 16
 Db 506 HKNKGNKNGHNGWKT 521

RESULT 2
 KGBOH2
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 22-Jun-1999
 C:Accession: A01282; A91923; A91941; A91938; B29559
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A:Title: A single gene for bovine high molecular weight and low molecular weight kin
 A:Reference number: A93317; PMID:84014106
 A:Accession: A01282
 A:Molecule type: mRNA
 A:Residues: 1-619 <KIT>
 A:Cross-references: GB:V01492; GB:K01758; NID:q493; PIDN:CAA24736.1; PID:q494
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
 A:Reference number: A91923; PMID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 376-391 <KAT>
 R:Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.

J. Biochem. 79, 1201-1222, 1976
 A>Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino acid sequence of the protein.
 A:Reference number: A91941; MUID:76260155
 A:Accession: A91941
 A:Molecule type: protein
 A:Residues: 387-455 <RAN>
 A>Note: 398-Pro, 401-Val, and 455-Lys were also found
 R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A>Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Amino acid sequence of the protein.
 A:Reference number: A91938; MUID:75170265
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 456-496 <RAN>
 R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga, S.
 J. Biol. Chem. 262, 2768-2779, 1987
 A>Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of cleavage sites, and the primary structure of the protein.
 A:Reference number: A92627; MUID:87137530
 A:Accession: B29559
 A:Molecule type: protein
 A:Residues: 12, 20-104, 106-256, 257-376 <SUE>
 R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A>Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen.
 A:Reference number: A91153; MUID:86030270
 A:Accession: A91153
 R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 808, 1984
 A>Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation: disulfide bonds
 A>Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as the LMW form is a cysteine protease inhibitor. Takes part in initiation of the C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is important for the release of bradykinin.
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-619/Product: HMW kininogen II #status predicted <MAT>
 F:19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
 F:19-130/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:261-372/Domain: cystatin homology <CY3>
 F:377-386/Product: lysyl-bradykinin (Kallidin II) #status experimental <KBDY>
 F:378-386/Product: bradykinin (Kallidin I) #status experimental <BDY>
 F:387-619/Product: HMW kininogen II light chain #status experimental <LCH>
 F:418-488/Region: glycine/histidine/lysine-rich
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
 F:27-589, 82-93, 106-125, 141-144, 205-217, 228-247, 261-264, 325-337, 348-367/Disulfide bonds:
 F:47/Binding site: carbonylate (Asn) (covalent) #status absent
 F:87, 168, 169, 204, 280/Binding site: carbonylate (Asn) (covalent) #status experimental
 F:136/Binding site: carbonylate (Thr) (covalent) (partial) #status experimental
 F:197/Binding site: carbonylate (Asn) (covalent) (partial) #status experimental
 F:376-377/Cleavage site: Met-Lys (Kallikrein) #status experimental
 F:380/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:386-387/Cleavage site: Arg-Ser (Kallikrein) #status experimental
 F:396, 400, 404, 510/Binding site: carbonylate (Ser) (covalent) #status experimental
 F:397, 398, 518, 522, 534, 546, 551, 568/Binding site: carbonylate (Thr) (covalent) #status experimental
 F:496-497/Cleavage site: Arg-Thr (Kallikrein) #status experimental

Query Match 76.5%; Score 75; DB 1; Length 619;
 Best Local Similarity 75.0%; Pred. No. 0.0019;
 Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 1 HKNKCKNGKNGKMT 16
 ||||| |||||
 Db 482 HKNKCKNGKNGKMT 497
 RESULT 3

KGB0H1
 A:Alternative names: alpha-2-thiol proteinase inhibitor; prekallikrein
 N:Contents: bradykinin (Kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A01281; A91923; A91938; A29559
 R:Kikuma, N.; Takagaki, Y.; Furuta, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A>Title: A single gene for bovine high molecular weight and low molecular weight kininogen.
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01281
 A:Molecule type: mRNA
 A:Cross-references: GB:V01491; GB:K01757; NID:q491; PIDN:CAA24735.1; PID:q492
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A>Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and the primary structure of the protein.
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 378-393 <KAT>
 R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A>Title: Studies on the primary structure of bovine high-molecular-weight kininogen.
 A:Reference number: A91938; MUID:75170265
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 458-498 <RAN>
 R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga, S.
 J. Biol. Chem. 262, 2768-2779, 1987
 A>Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of cleavage sites, and the primary structure of the protein.
 A:Reference number: A92627; MUID:87137530
 A:Accession: A29559
 A:Molecule type: protein
 A:Residues: 12, 20-123, 125-127, 129-378 <SUE>
 R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A>Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen.
 A:Reference number: A91153; MUID:86030270
 A:Accession: A91153
 R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 808, 1984
 A>Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation: disulfide bonds
 A>Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as the LMW form is a cysteine protease inhibitor. Takes part in initiation of the C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is important for the release of bradykinin.
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-621/Product: HMW kininogen I #status predicted <MAT>
 F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
 F:141-252/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:261-372/Domain: cystatin homology <CY3>
 F:377-386/Product: lysyl-bradykinin (Kallidin II) #status experimental <KBDY>
 F:380/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:386-387/Cleavage site: Arg-Ser (Kallikrein) #status experimental
 F:396, 400, 404, 510/Binding site: carbonylate (Ser) (covalent) #status experimental
 F:397, 398, 518, 522, 534, 546, 551, 568/Binding site: carbonylate (Thr) (covalent) #status experimental
 F:496-497/Cleavage site: Arg-Thr (Kallikrein) #status experimental

F:399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status ex
F:498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 76.5%; Score 75; DB 1; Length 621;
Best Local Similarity 75.0%; Pred. No. 0.0019;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 HKHKKKNGKHNGMKT 16
||||| ||||| |||||
DB 484 HKHKKKNGKHNGMKT 499

RESULT 4
T39083
conserved hypothetical protein SPAC7D4.05 - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39083

R:Gentles, S.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21826

A:Accession: T39083
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-225 <GEN>
A:Cross-references: EMBL:Z99532; PIDN:CAH16722.2; GSPDB:GN00066; SPDB:SPAC7D4.05

A:Experimental source: strain 97zh-; cosmid c7D4

C:Genetics:

A:Gene: SPDB:SPAC7D4.05

A:Map position: 1

A:Introns: 52/1; 77/3

C:Superfamily: Alkaligenes eutrophus phosphoglycolate phosphatase

Query Match 48.0%; Score 47; DB 2; Length 225;
Best Local Similarity 56.2%; Pred. No. 9.5;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

OY 1 HKHKKKNG--KHNGW 14
||| ||||| | | |
DB 60 HKHKKKSGLNPHDMW 75

RESULT 5
T35064
probable integral membrane protein - *Streptomyces coelicolor*

C:Species: *Streptomyces coelicolor*

C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C:Accession: T35064

R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A:Reference number: Z21567

A:Accession: T35064

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-275 <SEB>

A:Cross-references: EMBL:AL096884; PIDN:CAH51427.1; GSPDB:GN00070; SCOEDB:SC4G6.04C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC4G6.04C

C:Genetics:

A:Gene: SCOEDB:SC4G6.04C

C:Genetics:

A:Gene: SCOEDB:SC4G6.04C

C:Genetics:

A:Gene: SCOEDB:SC4G6.04C

S34153
mst101-1 protein - fruit fly (*Drosophila hydei*)

C:Species: *Drosophila hydei*

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000

C:Accession: S34153

R:Neesen, J.; Heinlein, U.A.O.; Buemann, H.

submitted to the EMBL Data Library, June 1993

A:Reference number: S34153

A:Accession: S34153

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-344 <NEE>

A:Cross-references: EMBL:X73480; NID:g313199; PID:g313200

C:Genetics:

A:Gene: FlyBase:Dhyd/mst101

A:Cross-references: FlyBase:FBgn0011816

C:Superfamily: neurofilament triplet H protein

Query Match 48.0%; Score 47; DB 2; Length 344;
Best Local Similarity 69.2%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 NKGGKNGKHNGWK 15
: ||||| | | |
DB 302 DKCKKNGKKNDMK 314

RESULT 7
G71333
probable ribonuclease H (rnhA) - *Syphilis spirochete*

C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)

C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999

C:Accession: G71333

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G

erson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; M

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A:Title: Complete genome sequence of *Treponema pallidum*, the *Syphilis spirochete*.

A:Reference number: A71250; MUID:98332770

A:Accession: G71333

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-169 <COL>

A:Cross-references: GB:AE001215; GB:AE000520; NID:g332631; PIDN:AA65340.1; PID:g332

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0353

C:Superfamily: ribonuclease H

C:Genetics:

A:Gene: TP0353

C:Genetics:

A:Gene: TP0353

C:Genetics:

A:Gene: TP0353

Query Match 46.9%; Score 46; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 KHNGWKT 16
||||| |
DB 103 KHNGWKT 109

RESULT 8
T49422
RAD57 related protein [imported] - *Neurospora crassa*

N:Alternate names: protein B17C10.30

C:Species: *Neurospora crassa*

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49422

R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49422

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-516 <SCH>
A:Cross-references: EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.30
A:Experimental source: BAC clone B17C10; strain OR74A
C:Genetics:
A:Gene: NCSP:B17C10.30
A:Map position: 6
A:Introns: 31/3

Query Match 46.9%; Score 46; DB 2; Length 516;
Best Local Similarity 64.3%; Pred. No. 28;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 KKKKKKKKKNGW 15
| | | | | | | | | | | | | | | | | |
DB 486 KGVKKKKKKKEGCK 499

RESULT 9

probable protein kinase (EC 2.7.1.-) ylop - Bacillus subtilis

C:Species: Bacillus subtilis
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: H69878

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Galitzzi, A.; Gall
lechl, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstapel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A.; Schreier, S.; Schroeter, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Serot
deuchl, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A65580; MUID:98044033

A:Accession: H69878

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-648 <KUN>

A:Cross-references: GB:Z99112; GB:AL009126; NID:92633902; PIDN:CAB13450.1; PID:92633949

A:Experimental source: strain 168

C:Genetics:

A:Gene: ylop

C:Superfamily: Bacillus subtilis probable protein kinase ylop; protein kinase homology
C:Keywords: ATP; phosphotransferase; protein kinase

F:9-269/Domain: protein kinase homology <KIN>

Query Match 46.9%; Score 46; DB 1; Length 648;
Best Local Similarity 61.5%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 KKKKKKKKKNGW 14
| | | | | | | | | | | | | | | | | |
DB 317 ENKTKKKKKRKKW 329

RESULT 10

hypothetical protein F8B4.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C:Accession: T05352

R:Bevan, M.; Terry, N.; Ardiles, W.; Buysnaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15409

A:Accession: T05352

A:Molecule type: DNA

A:Residues: 1-657 <BEV>

A:Cross-references: EMBL:AL0034567

A:Experimental source: cultivar Columbia; BAC clone F8B4

C:Genetics:

A:Map position: 4

A:Introns: 26/3; 45/1; 74/3; 83/1; 122/2; 165/1; 270/2; 307/1; 731/2; 754/2

A:Note: F8B4.120

C:Superfamily: cyclophilin homology
F:6-162/Domain: cyclophilin homology <CYP>

Query Match 46.9%; Score 46; DB 2; Length 857;
Best Local Similarity 80.0%; Pred. No. 44;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KKKKKKKKKH 11
| | | | | | | | | | | | | | | | | |
DB 168 KSDKKKKKKH 177

RESULT 11

thimet oligopeptidase (EC 3.4.24.15) pepb [validated] - Streptococcus agalactiae

C:Species: Streptococcus agalactiae

C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 15-Sep-2000

C:Accession: T51748

R:Lin, B.; Averett, W.F.; Novak, J.; Chatham, W.W.; Hollingshead, S.K.; Colligan, J.E.
Infect. Immun. 64, 3401-3406, 1996

A:Title: Characterization of PepB, a group B streptococcal oligopeptidase.

A:Reference number: Z25445; MUID:9633389

A:Accession: T51748

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-601 <LIN>

A:Cross-references: EMBL:U49821; PIDN:AAC44215.1

C:Genetics:

A:Gene: pepb

C:Function: A:Description: (EC 3.4.24.15) [validated; MUID:9633389]; hydrolyzes a variety of sma

C:Superfamily: oligopeptidase F

C:Keywords: hydrolase; metalloprotease

Query Match 45.9%; Score 45; DB 2; Length 601;
Best Local Similarity 53.8%; Pred. No. 45;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 HKKKKKKKKNG 13
| | | | | | | | | | | | | | | | | |
DB 349 HYNKKKRSATYSG 361

RESULT 12

hemoglobin V precursor - midge (Chironomus thummi piger)

C:Species: Chironomus thummi piger

C>Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 03-Mar-2000

C:Accession: S21633

R:Hankeln, T.; Rozynek, P.; Schmidt, E.R.
submitted to the EMBL Data Library, September 1990

A:Description: Complete nucleotide sequence of a hemoglobin gene cluster from the mid

A:Reference number: S21627

A:Accession: S21633

A:Molecule type: DNA

A:Residues: 1-163 <HAN>

A:Cross-references: EMBL:X56271; NID:97069; PIDN:CAA39718.1; PID:97076

C:Genetics:

A:Gene: hbv

A:Map position: II

C:Superfamily: globin; globin homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxygen carrier
F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-163/Product: hemoglobin V #status predicted <MAT>
F:19-162/Domain: globin homology <GLB>
F:76/Binding site: oxygen (His) (distal axial ligand) #status predicted
F:111/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 44.9%; Score 44; DB 1; Length 163;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 HKKGGKKNKGKKT 16
|||:|:|:|:|:|
Db 111 HKKGGITGGGNEKKT 126

RESULT 13
A:100
cell-division protein ftsX [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C:Accession: A84100

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; M01D:20263314
A:Accession: A84100

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-298 <STO>

A:Cross-references: GB:AP001519, GB:BA000004; NID:g10176109, PIDN:BA07320.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:

A:Gene: ftsX

Query Match 44.9%; Score 44; DB 2; Length 298;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 HKKGGKKNKGKKT 16
|:|:|:|:|:|
Db 9 HVREGTKNLGRNGWMT 24

RESULT 14

150680
alpha subunit of rod photoreceptor CNG-channel - chicken

C:Species: Gallus gallus (chicken)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Aug-1999

C:Accession: 150680

R:Boulyk, W.; Altenhofen, W.; Muller, F.; Dose, A.; Illing, M.; Molday, R.S.; Kaupp, U.B
Neuron 10, 865-877, 1993
A:Title: Rod and cone photoreceptor cells express distinct genes for cGMP-gated channels

A:Reference number: 150630; M01D:93264082

A:Accession: 150680

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-645 <BON>

A:Cross-references: EMBL:X89599; NID:g908852; PIDN:CA61758.1; PID:g908853

C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-
F:431-555/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>

Query Match 44.9%; Score 44; DB 2; Length 645;
Best Local Similarity 80.0%; Pred. No. 66;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 HKKGGKKNKGKKT 10
|||:|:|:|:|
Db 90 HKKGGKKNKGKKT 99

RESULT 15

A:86772

hypothetical protein yljD [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001

C:Accession: A86772

R:Bohlin, A.; Winkler, P.; Mager, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Eh
Genome Res. in press, 2001

A:Title: The complete genome sequence of the lactic acid bacterium.

A:Reference number: A86625

A:Accession: A86772

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-81 <STO>

A:Cross-references: GB:AE005176; NID:g12724143; PIDN:AAK05275.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yljD

Query Match 43.9%; Score 43; DB 2; Length 81;
Best Local Similarity 72.7%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 KKKGGKKNKGKKT 12
|:|:|:|:|:|
Db 35 KKKGGKKNKGKKT 45

Search completed: July 6, 2001, 09:18:03
Job time: 649 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:26:40 ; Search time 37.59 Seconds
(without alignments)
14.581 Million cell updates/sec

Title: US-09-437-912-9
Perfect score: 98
Sequence: 1 HKNGKKNGKNGKMT 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	75	76.5	619	1 KNH2_BOVIN	P01045 bos taurus
3	75	76.5	621	1 KNL_BOVIN	P01044 bos taurus
4	47	48.0	344	1 MSPL_DROHY	O08695 drephonema
5	46	46.9	169	1 RNH_TREPA	O83572 treponema p
6	45	45.9	601	1 PEPB_STRAG	Q53778 streptococc
7	44	44.9	163	1 GLBV_CHITP	P29243 chitronomus
8	44	44.9	645	1 CNG3_CHICK	O90980 gallus gall
9	43	43.9	174	1 RRS_PORPU	P51298 porphyra pu
10	43	43.9	374	1 YDVB_SCHPO	O14229 schizosacch
11	42.5	43.4	1532	1 IGA_NEIGO	P09790 neisseria g
12	42	42.9	678	1 GARP_PLAIF	P13816 plasmodium
13	41	41.8	186	1 PTR_HELIPJ	O92703 helicobacte
14	41	41.8	609	1 PEPF_BACSU	O31605 bacillus su
15	41	41.8	793	1 MURS_THEMA	P74926 thermotoga
16	41	41.8	812	1 TOP1_SCHPO	P07799 schizosacch
17	41	41.8	1372	1 INSR_MOUSE	P15208 mus musculu
18	41	41.8	1382	1 INSR_HUMAN	P06213 homo sapien
19	41	41.8	1383	1 INSR_RAT	P15127 rattus norv
20	40	40.8	99	1 RL23_HAEIN	P44561 haemophilus
21	40	40.8	102	1 RK23_ODOSI	P49559 odontella s
22	40	40.8	213	1 SKGR_XENLA	P13673 xenopus lae
23	40	40.8	304	1 RL2_AQUAE	O06434 aquifex aeo
24	40	40.8	312	1 YAO9_SCHPO	O10108 schizosacch
25	40	40.8	326	1 RNHL_SCHPO	O10236 schizosacch
26	40	40.8	421	1 SYTL_MOUSE	P46096 mus musculu
27	40	40.8	421	1 SYTL_RAT	P21707 rattus norv
28	40	40.8	422	1 SYTL_BOVIN	P48018 bos taurus
29	40	40.8	422	1 SYTL_HUMAN	P21579 homo sapien
30	40	40.8	424	1 SYTL_CHICK	P47191 gallus gall
31	40	40.8	427	1 SY61_DISOM	P24505 discospyge o
32	40	40.8	829	1 TOP1_XENLA	P45152 xenopus lae
33	40	40.8	861	1 TOP1_BUCAT	P57371 buchnera ap

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	644 AA.
1	KNG_HUMAN			
ID	P01042; P01043;			
AC	21-JUL-1986 (Rel. 01, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADIKININ].			
GN	KNG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).			
RC	TISSUE=Liver;			
RX	MEDLINE=85234582; PubMed=2989293;			
RA	Takagaki Y., Kitamura N., Nakanishi S.;			
RT	"Cloning and low molecular weight analysis of cDNAs for human high molecular weight and low molecular weight prekinnogens. Primary structures of two human prekinnogens.";			
RT	J. Biol. Chem. 260:8601-8609(1985).			
RN	[2]			
RP	GENE STRUCTURE.			
RX	MEDLINE=85234583; PubMed=2989294;			
RA	Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T., Nakanishi S.;			
RT	"Structural organization of the human kinnogen gene and a model for its evolution.";			
RT	J. Biol. Chem. 260:8610-8617(1985).			
RN	[3]			
RP	SEQUENCE OF 1-401 FROM N.A.			
RX	MEDLINE=8512621; PubMed=6441591;			
RA	Ohkubo I., Kurachi K., Takasawa T., Shiohara H., Sasaki M.;			
RT	"Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kinnogen.";			
RT	Biochemistry 23:5691-5697(1984).			
RN	[4]			
RP	SEQUENCE OF 379-644.			
RX	MEDLINE=8630270; PubMed=4054110;			
RA	Lottspeich F., Kellermann J., Henschen A., Foertsch B., Mueller-Esterl W.;			
RT	"The amino acid sequence of the light chain of human high-molecular-mass kinnogen.";			
RT	Eur. J. Biochem. 152:307-314(1985).			
RN	[5]			
RP	SEQUENCE OF 381-389.			
RX	MEDLINE=90255622; PubMed=4952632;			
RA	Pierce J.V.;			
RT	"Structural features of plasma kinins and kinnogens.";			
RT	Fed. Proc. 27:52-57(1968).			
RN	[6]			
RP	DISULFIDE BONDS.			
RA	Sueyoshi T., Miyata T., Kato H., Iwanaga S.;			
RT	"Disulfide bonds in bovine HMW kinnogens.";			

34	40	40.8	1203	1	SDC1_CAEEL	P24349 caenorhabdi
35	40	40.8	3255	1	POUG_LMYO	P31899 1 genome po
36	40	40.8	3255	1	POUG_LMYO	P86876 1 genome po
37	39	39.8	210	1	YEQ9_YEAST	P40052 saccharomyc
38	39	39.8	330	1	SYFA_CAMJE	O9P634 campylobact
39	39	39.8	420	1	DCDA_ECOLI	P00861 escherichia
40	39	39.8	491	1	NMT_CRYNE	P34809 cryptococcu
41	39	39.8	789	1	ACOX_YEAST	P39533 saccharomyc
42	39	39.8	874	1	POUL_HUMAN	P10266 homo sapien
43	39	39.8	1226	1	YCS3_YEAST	P25357 saccharomyc
44	39	39.8	1240	1	YNJ1_YEAST	P53935 saccharomyc
45	38.5	39.3	606	1	WDR1_HUMAN	O75083 homo sapien

RL Seikagaku 56:808-808(1984).

CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)

CC HMM-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY

CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO

CC FACTOR XII; (3) HMM-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-

CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE

CC BRADYKININ THAT IS RELEASED FROM HMM-KININOGEN SHOWS A VARIETY OF

CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE

CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND

CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A

CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR

CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF

CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS

CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,

CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)

CC LHM-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LHM-

CC KININOGEN IS IN CONTRAST TO HMM-KININOGEN NOT INVOLVED IN BLOOD

CC CLOTTING.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: HMM (SHOWN HERE) AND LHM; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: K02566; AAA35497.1; -

DR EMBL: M11437; AAB59550.1; -

DR EMBL: M11438; AAB59550.1; JOINED.

DR EMBL: M11521; AAB59550.1; JOINED.

DR EMBL: M11522; AAB59550.1; JOINED.

DR EMBL: M11523; AAB59550.1; JOINED.

DR EMBL: M11524; AAB59550.1; JOINED.

DR EMBL: M11525; AAB59550.1; JOINED.

DR EMBL: M11526; AAB59550.1; JOINED.

DR EMBL: M11527; AAB59550.1; JOINED.

DR EMBL: M11528; AAB59550.1; JOINED.

DR EMBL: M11437; AAB59551.1; -

DR EMBL: M11438; AAB59551.1; JOINED.

DR EMBL: M11521; AAB59551.1; JOINED.

DR EMBL: M11523; AAB59551.1; JOINED.

DR EMBL: M11524; AAB59551.1; JOINED.

DR EMBL: M11525; AAB59551.1; JOINED.

DR EMBL: M11526; AAB59551.1; JOINED.

DR EMBL: M11527; AAB59551.1; JOINED.

DR EMBL: M11528; AAB59551.1; JOINED.

DR PIR: A01279; KGH0H1.

DR PIR: A25276; A25276.

DR PIR: A01280; KGH0L1.

DR PIR: B25276; B25276.

DR PIR: S02482; S02482.

DR SWISS-2DPAGE: P01043; HUMAN.

DR MIM: 228960; -

DR InterPro: IPR000010; -

DR InterPro: IPR002395; -

DR Pfam: PF00031; CYSTATIN_3.

DR PRINTS: PR00334; KININOGEN.

DR PROSITE: PS00287; CYSTATIN_2.

KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;

KW Bradykinin; Blood coagulation; Inflammatory response; Signal;

KW Alternative splicing.

FT SIGNAL 1 18

FT CHAIN 19 644 KININOGEN,

FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 BRADYKININ.

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD_RES	19	19	
FT	DISULFID	28	614	PYRROLIDONE CARBOXYLIC ACID.
FT	DISULFID	83	94	INTERCHAIN.
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	205	205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	294	294	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPPLIC	402	427	
FT	VARSPPLIC	428	644	VSPPTSNAPADEERDSKGEQGHTR -> SHLRSCYKGR
FT	CONFLICT	593	593	PKRGAEPASREKRS (IN ISOFORM LHM).
FT	SEQUENCE	644 AA:	71945 MM:	MISSING (IN ISOFORM LHM).
FT	SEQUENCE			T -> I (IN REF. 1).
FT	SEQUENCE			3132B4CBAF8FB7E CRC64;

Query Match 100.0%; Score 98; DB 1; Length 644;

Best Local Similarity 100.0%; Pred. NO. 7.3e-08;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKNKGRKNGKNGKWT 16

Db 506 HKNKGRKNGKNGKWT 521

RESULT 2

ID KNH2_BOVIN STANDARD; PRT; 619 AA.

AC P01045;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE KININOGEN, HMM II PRECURSOR (THIOLE PROTEINASE INHIBITOR) [CONTAINS: DE BRADYKININ].

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

OX [1]

RP SEQUENCE FROM N.A.

RP MEDLINE=84014106; PubMed=6571699;

RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nakanishi S.;

RT "A single gene for bovine high molecular weight and low molecular weight kininogens."

RL Nature 305:545-549(1983).

RN [2]

RP SEQUENCE OF 19-376.

RP MEDLINE=87137530; PubMed=3546295;

RX Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.;

RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion.";
 RT J. Biol. Chem. 262:2768-2779(1987).
 RN [3]
 RP SEQUENCE OF 376-391.
 RX MEDLINE-70180420; PubMed-4986212;
 RA Kato H., Nagasawa S., Suzuki T.;
 RT Studies on the structure of bovine kininogen: cleavages of disulfide
 RT bonds and of methionyl bonds in kininogen-II.";
 RT J. Biochem. 67:313-323(1970).
 RN [4]
 RP SEQUENCE OF 387-455.
 RX MEDLINE-76260155; PubMed-956151;
 RA Han Y.N., Kato H., Iwanaga S., Suzuki T.;
 RT "Primary structure of bovine plasma high-molecular-weight kininogen.
 RT The amino acid sequence of a glycopeptide portion (fragment 1)
 RT following the C-terminus of the bradykinin moiety.";
 RT J. Biochem. 79:1201-1222(1976).
 RN [5]
 RP SEQUENCE OF 456-496.
 RX MEDLINE-75170265; PubMed-1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 RT kininogen. Amino acid sequence of a fragment ('histidine-rich
 RT peptide') released by plasma kallikrein.";
 RT J. Biochem. 77:55-68(1975).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C)
 CC NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4E4) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 398.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: V01492; CAA24736.1; -
 CC EMBL: V01492; CAA24737.1; ALT_SEQ.
 CC PIR: A01282; KGBH2.
 CC PIR: B29559; B29559.
 CC HSSP: P04129; IAF1.
 CC InterPro: IPR000010; -
 CC InterPro: IPR002395; -
 CC Pfam: PF00031; cystatin. 3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSTATIN. 2.
 CC Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 CC Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
 CC Inflammatory response.

FT SIGNAL 1 18
 FT CHAIN 19 619
 FT CHAIN 19 376
 FT CHAIN 19 376
 FT PEPTIDE 378 386
 FT CHAIN 387 619
 FT DOMAIN 19 135
 FT DOMAIN 136 256
 FT DOMAIN 257 376
 FT MOD.RES 19 19
 FT CARBOHYD 87 87
 FT CARBOHYD 136 136
 FT CARBOHYD 168 168
 FT CARBOHYD 197 197
 FT CARBOHYD 204 204
 FT CARBOHYD 280 280
 FT CARBOHYD 400 400
 FT DISULFID 27 589
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 261 264
 FT DISULFID 325 337
 FT DISULFID 348 367
 FT VARIANT 398 398
 FT VARIANT 401 401
 FT VARIANT 454 454
 SQ SEQUENCE 619 AA; 68710 MW; F04320A8E0E0DA CRC64;
 Query Match 76.5%; Score 75; DB 1; Length 619;
 Best Local Similarity 75.0%; Pred. No. 0.00025;
 Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 HKNKGRKNGKNGKMT 16
 DB 482 HKNKGRKNGKNGKMT 497
 RESULT 3
 ID KNL1_BOVIN STANDARD; PRT; 621 AA.
 AC P01044;
 DT 21-JUL-1986 (rel. 01, Created)
 DT 21-JUL-1986 (rel. 01, Last sequence update)
 DT 01-JUN-1994 (rel. 29, Last annotation update)
 DE KININOGEN, HMW I PRECURSOR (THIOLE PROTEINASE INHIBITOR) [CONTAINS:
 DE BRADYKININ].
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-84014106; PubMed-6571699;
 RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
 RT "A single gene for bovine high molecular weight and low molecular
 RT weight kininogens.";
 RT Nature 305:545-549(1983).
 RN [2]
 RP SEQUENCE OF 19-378.
 RX MEDLINE-87137350; PubMed-3546295;
 RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
 RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion.";
 RT J. Biol. Chem. 262:2768-2779(1987).
 RN [3]
 RP SEQUENCE OF 378-393.
 RX MEDLINE-70180420; PubMed-4986212;

RA Kato H., Nagasawa S., Suzuki T.:
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [4]
 RP SEQUENCE OF 458-498.
 RX MEDLINE=75170265; PubMed=1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.:
 RT "Studies on the primary structure of bovine high-molecular-weight
 kininogen. Amino acid sequence of a fragment ('histidine-rich
 peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
 CC NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 400.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: V01491; CAA24735.1; -
 CC PIR: A01281; KGB0H1.
 CC PIR: A29559; A29559.
 CC InterPro: IPR000010; -
 CC InterPro: IPR002395; -
 CC Pfam: PF00031; cystatin; 3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSTATIN; 2.
 CC Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 CC Thiol protease inhibitor; Bradykinin; Blood coagulation;
 CC Inflammatory response; Signal.
 CC
 CC SIGNAL 1 18
 CC FT CHAIN 19 621 KININOGEN, HMW I.
 CC FT CHAIN 19 378 HEAVY CHAIN.
 CC FT PEPTIDE 380 388 BRADYKININ.
 CC FT CHAIN 389 621 LIGHT CHAIN.
 CC FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 CC FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 CC FT DOMAIN 258 378 CYSTATIN-LIKE 3.
 CC FT MOD. RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 CC FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 CC FT CARBOHYD 136 136 PARTIAL.
 CC FT CARBOHYD 168 168 OR 169.
 CC FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 CC FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 CC FT DISULFID 27 591 INTERCHAIN.
 CC FT DISULFID 82 93
 CC FT DISULFID 106 125
 CC FT DISULFID 141 144

FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 FT DISULFID 621 6890 MM; D16850BEFE3C55CD CRC64;
 SQ SEQUENCE

Query Match 76.5%; Score 75; DB 1; Length 621;
 Best local Similarity 75.0%; Pred. No. 0 00025;
 Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 HKNKKKKNGKNGKWT 16
 DB 484 HKNKKKKNGKNGKWDRT 499

RESULT 4
 ID MST1_DROHY STANDARD; PRT; 344 AA.
 AC 008695;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE AXONEME-ASSOCIATED PROTEIN MST101(1).
 GN MST101(1).
 OS Drosophila hydei (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7224;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE-Testis;
 RX MEDLINE=94200512; PubMed=8150205;

RA Neesen J., Buemann H., Heinlein U.A.:
 RT "The Drosophila hydei gene Dhmst101(1) encodes a testis-specific,
 RT repetitive, axoneme-associated protein with differential abundance in
 RT x chromosome deletion mutant flies";
 RL Dev. Biol. 162:414-425(1994).
 CC -1- FUNCTION: POSSIBLE STRUCTURAL ROLE IN THE SPERM TAIL. IT IS
 CC ASSOCIATED WITH AXONEMAL STRUCTURES.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: TESTIS. LOCATED IN SPERMATOCTES AND
 CC SPERMATID BUNDLES.
 CC -1- DOMAIN: THE PREDOMINANT STRUCTURE IS ALPHA-HELICAL.
 CC -1- POLYMORPHISM: THE NUMBER OF REPEATS VARIES BETWEEN STRAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X73480; CAA51875.1; -
 CC PIR: S34153; S34153.
 CC DR HSSP: P01032; IC5A.
 CC DR FlyBase: FBgn0011816; Dhyd\mst101(1).
 CC Sperm; Repeat; Multigene family.
 CC DOMAIN 58 337 19 X 16 AA APPROXIMATE TANDEM REPEATS OF
 CC K-K-K-C-X-E-X-A-[KQ]-K-X-X-E-X-A-X.

SO SEQUENCE 344 AA; 37793 MM; 2465D2510387E2A CRC64;

Query Match 48.0%; Score 47; DB 1; Length 344;
 Best local Similarity 69.2%; Pred. No. 3.1;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 MKGKKNGKNGKWK 15
 DB 302 DKGKKNGKNGKDKM 314

RESULT 5
ID RNH_TREPA STANDARD; PRT; 169 AA.
AC 083372;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBONUCLEASE H (EC 3.1.26.4) (RNASE H).
GN RNHA OR TP0353.
OS Treponema pallidum.
OC Bacteri: Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxId=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Ullrich T.,
RA McDonald L., Atliach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete.";
RL Science 281:375-388(1998).
CC -1- FUNCTION: THIS ENZYME IS AN ENDONUCLEASE THAT DEGRADES THE RNA OF
CC RNA-DNA HYBRIDS SPECIFICALLY (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
CC MONESTER.
CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE RNASE H FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001215; AAC65340.1; -
DR TIGR: TP0353; -
DR InterPro: IPR002156; -
DR Pfam: PF00075; rnaaseh.1.
KW Hydrolase; Nuclease; Endonuclease; Magnesium.
FT METAL 12 12 MAGNESIUM (BY SIMILARITY).
FT METAL 63 63 MAGNESIUM (BY SIMILARITY).
FT METAL 87 87 MAGNESIUM (BY SIMILARITY).
FT METAL 151 151 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 169 AA; 18184 MW; 164311053632B047 CRC64;

Query Match 46.9%; Score 46; DB 1; Length 169;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KHNKMK 16
ID 11111111
DB 103 KHNKMK 109

RESULT 6
PEPB_STRAG STANDARD; PRT; 601 AA.
AC 053778;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE GROUP B OLIGOPEPTIDASE PEPB (EC 3.4.24.-).
GN PEPB.
OS Streptococcus agalactiae.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxId=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3502;
RX MEDLINE=9633389; PubMed=8757883;
RA Lin B., Averett W.F., Novak J., Chatham W.W., Hollingshead S.K.,
RA Coligan J.E., Egan M.L., Pritchard D.G.;
RT "Characterization of PEPB, a group B streptococcal oligopeptidase."
RL Infect. Immun. 64:3401-3406(1996).
CC -1- FUNCTION: HAS OLIGOPEPTIDASE ACTIVITY AND DEGRADES A VARIETY OF
CC SMALL BIOACTIVE PEPTIDES, INCLUDING BRADIKININ, NEUTROSTENIN, AND
CC PEPTIDE FRAGMENTS OF SUBSTANCE P AND ADRENOCORTICOTROPIN. ALSO
CC HYDROLYZES THE SYNTHETIC COLLAGEN-LIKE SUBSTRATE N-(3-[2-
CC FURYL]ACRYLOYL)-LEU-GLY-PRO-ALA (FALGPA).
CC -1- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M3 (ZINC METALLOPROTEASE);
CC ALSO KNOWN AS THE THIMET OLIGOPEPTIDASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U49821; AAC44215.1; -
DR MEROPS: M03.008; -
DR InterPro: IPR00130; -
DR InterPro: IPR001567; -
DR Pfam: PF01432; Peptidase_M3.1.
DR PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Metalloprotease; Hydrolase; Zinc.
FT METAL 386 386 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 387 387 BY SIMILARITY.
FT METAL 390 390 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 393 393 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 601 AA; 69590 MW; EE524998D30B08F CRC64;

Query Match 45.9%; Score 45; DB 1; Length 601;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 HKNKKNKNGKNG 13
ID 11111111
DB 349 HKNKGRKSGAYSG 361

RESULT 7
GLBV_CHITP
ID GLBV_CHITP STANDARD; PRT; 163 AA.
AC P29243;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GLOBIN CTT-V PRECURSOR (HBV).
GN CTT-V.
OS Chironomus thummi piger (Midge).
OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta;
OC Pterygota: Neoptera: Endopterygota: Diptera: Nematocera;
OC Chironomidae: Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxId=7156;
RN [1]
RP SEQUENCE FROM N.A.
RA Hakein T., Rozynek P., Schmidt E.R., Broecker M.;
RL Submitted (SEP-1990) to the EMBL/GenBank/DBJ databases.

```

CC -1- MISCELLANEOUS: THERE ARE AT LEAST 12 DIFFERENT COMPONENTS IN MIDGE
CC GLOBIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X56271; CAA39718.1; -.
CC DR PIR: S21633; S21633.
CC DR HSSP: P02229; IECO.
CC DR InterPro: IPR000971; -.
CC DR InterPro: IPR002336; -.
CC DR Pfam: PF00042; globin.1.
CC DR PRINTS: PRO0611; ERYTHROGRIN.
CC DR PROSITE: PS01033; GLOBIN; 1.
CC KW Heme; Oxygen transport; Respiratory protein; Signal.
CC FT SIGNAL 1 16 GLOBIN CTT-V.
CC FT CHAIN 17 163 IRON (HEME DISTAL LIGAND)
CC FT METAL 76 76 (BY SIMILARITY).
CC FT METAL 111 111 IRON (HEME PROXIMAL LIGAND)
CC FT METAL (BY SIMILARITY).
CC SQ SEQUENCE 163 AA; 17782 MW; 8E8567BC630A646E CRC64;

Query Match 44.98; Score 44; DB 1; Length 163;
Best Local Similarity 50.0%; Pred. No. 4.3;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 HKKNGKKNGKNGMKT 16
   |||:| 1:1:1|
Db 111 HKKNGITGGFNEFKT 126

RESULT 8
CN3-CHICK STANDARD; PRT; 645 AA.
AC 090980;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CYCLIC NUCLEOTIDE GATED CHANNEL, ROD PHOTORECEPTOR, ALPHA SUBUNIT
DE (CNG CHANNEL 3) (CNG-3) (CNG3).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93264082; PubMed-7684234;
RA Boenigk W., Alenmofen W., Mueller F., Dose A., Illing M.,
RA Molday R.S., Kaupp U.B.;
RA "Rod and cone photoreceptor cells express distinct genes for
RA cGMP-gated channels.";
RT Neuron 10:865-877(1993).
RL
CC -1- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN
CC COUPLED CASCADE USING GMP AS SECOND MESSENGER. THIS PROTEIN CAN
CC BE ACTIVATED BY CYCLIC GMP WHICH LEADS TO A OPENING OF THE CATION
CC CHANNEL AND THEREBY CAUSING A DEPOLARIZATION OF ROD
CC PHOTORECEPTORS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X89599; CAA61758.1; -.
CC DR InterPro: IPR000595; -.
CC DR InterPro: IPR002025; -.
CC DR Pfam: PF00914; CNG_membrane; 1.
CC DR Pfam: PF00027; CNMP_binding; 1.
CC DR PROSITE: PS00888; CNMP_BINDING_1; 1.
CC DR PROSITE: PS00889; CNMP_BINDING_2; 1.
CC DR PROSITE: PS50042; CNMP_BINDING_3; 1.
CC KW Ionic channel; Ion transport; cAMP-binding; Transmembrane; Vision;
CC Multigene family.
CC FT DOMAIN 1 117 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 118 137 H1 (POTENTIAL).
CC FT DOMAIN 138 150 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 151 169 H2 (POTENTIAL).
CC FT DOMAIN 170 193 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 194 213 H3 (POTENTIAL).
CC FT DOMAIN 214 251 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 252 274 H4 (POTENTIAL).
CC FT DOMAIN 275 326 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 327 346 H5 (POTENTIAL).
CC FT DOMAIN 347 430 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 431 451 H6 (POTENTIAL).
CC FT DOMAIN 452 645 CYTOPLASMIC (POTENTIAL).
CC FT NP_BIND 439 561 CAMP (BY SIMILARITY).
CC FT BINDING 498 498 CAMP (POTENTIAL).
CC FT BINDING 513 513 CAMP (POTENTIAL).
CC FT CARBOHYD 377 377 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SQ SEQUENCE 645 AA; 74778 MW; 989D515F61ACTD31 CRC64;

Query Match 44.98; Score 44; DB 1; Length 645;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 HKKNGKKNGK 10
   |||||
Db 90 HKKNDKKKK 99

RESULT 9
RR5-PORPU STANDARD; PRT; 174 AA.
AC P51298;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S5.
DE REPS.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-AVONPORT;
RA Reith M.E., Munholland J.;
RA "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RA genome.";
RT Plant Mol. Biol. Rep. 13:333-335(1995).
RL
CC -1- SIMILARITY: BELONGS TO THE S5P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR EMBL: U38804; AAC08184.1; -
DR HSSP: P02357; 1PKP.
DR Mendel; 10355; PORPU: rps5.1.
DR InterPro: IPR000851; -
DR Pfam: PF00333; Ribosomal_S5; 1.
DR PROSITE: PS00385; RIBOSOMAL_S5; 1.
DR RIBOSOMAL protein; Chloroplast.
KW SEQUENCE 174 AA; 18294 MW; 9052BDA33262460C CRC64;

Query Match 43.9%; Score 43; DB 1; Length 174;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 KKKKKKKHNGWK 15
| : | | | | |
DB 5 KKKKKKKHNGWK 18

RESULT 10
YDVB_SCHPO STANDARD; PRT; 374 AA.
ID YDVB_SCHPO
AC 014229;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOHERETICAL 43.4 KDA PROTEIN C6F12.11C IN CHROMOSOME I.
GN SPAC6F12.11C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z98533; CAB11095.1; -
CC KW Hypothetical protein.
CC FT DOMAIN 66 69 POLY-PRO
SQ SEQUENCE 374 AA; 43426 MW; 615C054F6AE0D471 CRC64;

Query Match 43.9%; Score 43; DB 1; Length 374;
Best Local Similarity 69.2%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 KKKKKKKHNGWK 16
| : | | | | |
DB 212 KKKKKKKHNGWK 224

RESULT 11
IGA_NEIGO STANDARD; PRT; 1532 AA.
ID IGA_NEIGO
AC P09790;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE IGA-SPECIFIC SERINE ENDOPEPTIDASE PRECURSOR (EC 3.4.21.72) (IGA
DE PROTEASE).
GN IGA.
OS Neisseria gonorrhoeae.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=MS11;
RX MEDLINE=87115823; PubMed=3027577;
RA Pohlner J., Hailer R., Beyreuther K., Meyer T.F.;
RT "gene structure and extracellular secretion of Neisseria gonorrhoeae
RT IGA protease.";
RL Nature 325:458-462(1987).

RN [2]
RX MEDLINE=90154052; PubMed=2105953;
RA Bachovchin W.M., Plant A.G., Flentke G.R., Lynch M., Retner C.A.;
RT "Inhibition of Iga1 proteinases from Neisseria gonorrhoeae and
RT Hemophilus influenzae by peptide prolyl boronic acids.";
RL J. Biol. Chem. 265:3738-3743(1990).
CC -1- FUNCTION: THIS PROTEASE IS SPECIFIC FOR IMMUNOGLOBULIN A.
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 56 (SERINE PROTEASE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X04835; CAA28538.1; -
DR PIR: A26039; A26039.
DR MEROPS: S06.001;
DR InterPro: IPR000710; -
DR PRINTS: PRO00921; IGASERPRASE.
KW Hydrolyase; Serine protease; Zymogen; Autocatalytic cleavage;
KW Transmembrane; Signal.
FT SIGNAL 1 27
FT CHAIN 28 986 IGA-SPECIFIC SERINE ENDOPEPTIDASE.
FT PROPEP 987 1532 HELPER PEPTIDE.
FT ACT_SITE 278 278 POTENTIAL.
FT SITE 986 987 CLEAVAGE (AUTO-).
FT SITE 1018 1019 CLEAVAGE (AUTO-).
FT SITE 1121 1122 CLEAVAGE (AUTO-).
SQ SEQUENCE 1532 AA; 168976 MW; 68F4112BD22F40D CRC64;

Query Match 43.4%; Score 42.5; DB 1; Length 1532;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 2;

OY 1 HKNKG--KKKKHNGWK 16
| : | | | | |
DB 326 HDNAGTVKNGEHH-WKT 342

RESULT 12
GARP_PLAFF STANDARD; PRT; 678 AA.
ID GARP_PLAFF
AC P13816;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLUTAMIC ACID-RICH PROTEIN PRECURSOR.
GN GARP.
OS Plasmodium falciparum (Isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89040048; PubMed=2903445;
RA Triglia T., Stehl H.-D., Crekter P.E., Silva A., Anders R.F.,
Kemp D.O.;
RT "Structure of a Plasmodium falciparum gene that encodes a glutamic
acid-rich protein (GARP).";
RL Mol. Biochem. Parasitol. 31:199-202(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03998; AAA29605.1; -
DR PIR: A54514; A54514.
KW Repeat; Malaria; Antigen; Signal.
FT SIGNAL 1 25
FT CHAIN 26 678
FT DOMAIN 120 164 15 X 3 AA TANDDEM REPEATS OF K-K-X.
FT DOMAIN 372 416 9 X APPROXIMATE TANDDEM REPEATS.
FT DOMAIN 417 441 5 X APPROXIMATE TANDDEM REPEATS.
FT DOMAIN 576 604 POLY-GLU.
FT DOMAIN 605 653 7 X APPROXIMATE TANDDEM REPEATS.
FT DOMAIN 654 663 POLY-GLU
SQ SEQUENCE 678 AA: 80551 MW: 2A8F85606496EA9E CRC64;

Query Match 42.9%; Score 42; DB 1; Length 678;
Best Local Similarity 52.9%; Pred. No. 35;
Matches 9; Conservative 2; Mismatches 0; Indels 6; Gaps 1;

QY 1 HKKNGKKN-----GKH 11
DB 414 HKSKGKKDKGKDKRKH 430

RESULT 13
PTH_HELPJ STANDARD; PRT; 186 AA.
AC 09ZJC3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PEPTIDYL-TRNA HYDROLASE (EC 3.1.1.29) (PTH).
GN PTH OR JHP1390.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Molt D.T., King B.L., Brown E.D., Dolg P.C.,
Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -----
CC FUNCTION: THE NATURAL SUBSTRATE FOR THIS ENZYME MAY BE PEPTIDYL-
CC TRNS WHICH DROP OFF THE RIBOSOME DURING PROTEIN SYNTHESIS
CC (BY SIMILARITY).
CC CATALYTIC ACTIVITY: N-SUBSTITUTED AMINOACYL-TRNA + H(2)O =
CC N-SUBSTITUTED AMINO ACID + TRNA.
CC SUBUNIT: MONOMER (BY SIMILARITY).
CC SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC SIMILARITY: BELONGS TO THE PTH FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: A601561; AAD06971.1; -
DR HSSP; P23932; 2PTH.
DR InterPro: IPR001328; -
DR Pfam: PF01195; Pept_trna_hydro. 1.
DR PROSITE: PS01195; PEPT_TRNA_HYDROL_1; 1.
DR PROSITE: PS01196; PEPT_TRNA_HYDROL_2; 1.
KW Hydrolase.
SQ SEQUENCE 186 AA: 20956 MW: 95315FDF43D06112 CRC64;

Query Match 41.8%; Score 41; DB 1; Length 186;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 GKKNKGKNGKWT 16
DB 99 KGGNGGNGHGLKS 111

RESULT 14
PEPT_BACSU STANDARD; PRT; 609 AA.
ID PEPT_BACSU
AC 031605;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE OLIGOENDOPEPTIDASE F HOMOLOG (EC 3.4.24.-).
GN YJBG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Dancho A.;
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
CC CORYCTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY M3 (ZINC METALLOPROTEASE);
CC ALSO KNOWN AS THE THIMET OLIGOPEPTIDASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z99110; CAB13011.1; -
DR Subtilist; BG1136; yjbg.
DR InterPro: IPR001530; -
DR InterPro: IPR001567; -
DR Pfam: PF01432; Peptidase_M3; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc.
FT METAL 395 395 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 396 396 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 399 399 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 609 AA: 70144 MW: 674B0AFD6FA0D075 CRC64;

Query Match 41.8%; Score 41; DB 1; Length 609;

Best Local Similarity 50.0%; Pred. No. 45;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
OY 1 HKKGRKNGKH 12
Db 358 YENKGRNGAYS 369

OY 2 KKKGRKNGKH 12
Db 772 KNTKKNGKSN 782

Search completed: July 6, 2001, 09:26:40
Job time: 971 sec

RESULT 15
MUTS_THEME
ID MUTS_THEME STANDARD; PRT; 793 AA.
AC P74926;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA MISMATCH REPAIR PROTEIN MUTS.
GN MUTS OR TM1719.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID-2336;
RN [1]
RP SEQUENCE FROM N.A.
RA Wetmur J.G., Rosenfeld A., Wong D.M.;
RT "Hyperthermophilic Muts proteins: Isolation, characterization and
enhancement of PCR specificity."
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA STRAIN-MSB8 / DSM 3109;
RX MEDLINE-99287316; PubMed-10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwyn M.L., Dodson R.J.,
Hafit D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RA "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima."
RT Nature 399:323-329 (1999).
RL Nature 399:323-329 (1999).
CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES
IN DNA. IT IS POSSIBLE THAT IT CARRY OUT THE MISMATCH RECOGNITION
STEP. THIS PROTEIN HAS A WEAK ATPASE ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U71155; AAB16999.1; -
DR EMBL: AE001811; AAD36785.1; -
DR TIGR: TM1719; -
DR InterPro: IPR000432; -
DR InterPro: IPR002863; -
DR Pfam: PF00488; MUTS_C; 1.
DR Pfam: PF01624; MUTS_N; 1.
DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
KW DNA repair; ATP-binding; DNA-binding.
FT NP_BIND 589 596 ATP (POTENTIAL).
FT CONFLICT 233 233 A -> G (IN REF. 1).
FT CONFLICT 262 262 L -> W (IN REF. 1).
FT CONFLICT 287 287 L -> G (IN REF. 1).
FT CONFLICT 506 506 K -> T (IN REF. 1).
SQ SEQUENCE 793 AA; 91065 MW; 1BCB2342E4F9B1BD CRC64;

Query Match 41.8%; Score 41; DB 1; Length 793;
Best Local Similarity 72.7%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

THIS PAGE BLANK (unrot)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:25:56 ; Search time 118.42 Seconds

(without alignments)
17.876 Million cell updates/sec

Title: US-09-437-912-9

Perfect score: 98

Sequence: 1 HKNKGKNGKNGKNGKWT 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-unclassified:*
- 13: sp-vertebrate:*
- 14: sp-virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	51.0	581	10	Q9SEL2 vitis vinif
2	49	50.0	3898	14	Q68965 classical s
3	47	48.0	225	3	014262 schizosacch
4	47	48.0	275	2	09S2U7 streptomyc
5	46	46.9	95	1	Q9HL50 thermoplasm
6	46	46.9	516	3	Q9P6E6 bacillus su
7	46	46.9	648	2	034507 neurospora
8	46	46.9	857	10	Q9SUVO arabidopsis
9	46	46.9	1372	4	Q9HAW1 homo sapien
10	46	46.9	1388	4	Q9HAW2 homo sapien
11	46	46.9	2187	4	Q9H197 homo sapien
12	46	46.9	2254	4	Q9HCY0 homo sapien
13	45	45.9	1251	5	016337 caenorhabd
14	44	44.9	298	2	Q9K6X3 bacillus ha
15	44	44.9	806	10	Q9FIM2 schizosacch
16	43	43.9	171	3	Q9UTM3 plasmodium
17	43	43.9	504	5	097467 arabidopsis
18	43	43.9	808	10	023052 arabidopsis
19	43	43.9	915	5	Q9GCU7 plasmodium

20	42.5	43.4	205	4	Q9P0Q2 homo sapien
21	42.5	43.4	215	4	Q9NMU9 homo sapien
22	42.5	43.4	427	4	Q9Y5A8 homo sapien
23	42.5	43.4	823	5	Q9VFR5 drosophila
24	42.5	43.4	953	11	Q9RIC7 mus musculus
25	42	42.9	83	5	Q23783 chironomus
26	42	42.9	111	5	Q9W2P6 drosophila
27	42	42.9	205	14	Q9DPT6 human immun
28	42	42.9	240	5	Q23780 chironomus
29	42	42.9	294	4	Q9Y521 homo sapien
30	42	42.9	673	5	Q9U0N1 plasmodium
31	42	42.9	682	5	Q9GTM3 caenorhabd
32	42	42.9	709	5	Q16789 caenorhabd
33	42	42.9	739	2	Q87381 haemophilus
34	42	42.9	746	10	Q49546 arabidopsis
35	42	42.9	749	5	Q17768 caenorhabd
36	42	42.9	800	4	Q9N7L2 homo sapien
37	42	42.9	800	4	Q9NVE9 homo sapien
38	42	42.9	800	4	Q9HDD2 homo sapien
39	42	42.9	800	4	Q9HCC5 homo sapien
40	42	42.9	842	10	Q9LUN5 ratius norv
41	42	42.9	1009	11	Q08764 arabidopsis
42	42	42.9	1230	10	Q9SZY0 xiphophorus
43	42	42.9	1503	13	Q918X6 classical s
44	42	42.9	3898	14	Q68534 classical s
45	42	42.9	3898	14	Q92364 classical s

ALIGNMENTS

RESULT 1: PRELIMINARY: PRT; 581 AA.

Q9SEL2: 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-MAY-2001 (TREMREL. 16, Last annotation update)
DE GAG-POL. POLYPROTEIN.
OS Vitis vinifera (Grape).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Vitaceae; Vitis.
OX NCBI_TaxID=29760;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DANUTA; TISSUE=YOUNG LEAVES;
RA Vertes C.; Bes C.; Teshiere C.;
RT "VINE 1: a long terminal repeat element integrated in an Adh gene of
cv. Danuta and dispersed in this Vitis vinifera genome."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF16598; AAF20282.1; -
DR InterPro; IPR000209; -
DR InterPro; IPR001584; -
DR InterPro; IPR001878; -
DR Pfam; PF00665; rve; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
DR SMART; SM00343; ZnF_C2HC; 1.
KW Polyprotein.
SQ SEQUENCE 581 AA; 66488 MW; A908822FFB09B41D CRC64;

Query Match 51.0%; Score 50; DB 10; Length 581;
Best Local Similarity 62.5%; Pred. No. 5.3;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 HKNKGKNGKNGKNGKWT 16
Db 218 HKNKGKNGKNGKNGK 233

RESULT 2
Q68965
ID Q68965 PRELIMINARY: PRT; 3898 AA.

AC 068965;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE HOG CHOLERA VIRUS POLYPROTEIN.
 OS Classical swine fever virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Pestivirus.
 OX NCBI_TaxID=11096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GPE-;
 AC MEDLINE=95390794; PubMed-7661692;
 RX Ishikawa K., Nagai H., Katayama K., Tsutsui M., Tanabayashi K.,
 RA Takeuchi K., Hishiyama M., Saitoh A., Takagi M., Gotoh K.,
 RA Muramatsu M., Yamada A.;
 RT "Comparison of the entire nucleotide and deduced amino acid sequences
 of the attenuated hog cholera vaccine strain GPE- and the wild-type
 parental strain A.D.".
 RL Arch. Virol. 140:1385-1391(1995).
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 DR EMBL: D49533; BAA08477.1; -;
 DR InterPro: IPR000280; -;
 DR InterPro: IPR001005; -;
 DR InterPro: IPR001410; -;
 DR InterPro: IPR001568; -;
 DR InterPro: IPR001650; -;
 DR Pfam: PF00271; helicase C; 1.
 DR PRINTS: PR00729; CDVENDOPTASE.
 DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
 DR PROSITE: PS00531; RNASE_T2.2; UNKNOWN_1.
 DR SMART: SM00490; HELIC; 1.
 KW ATP-binding; Helicase; Polypeptide.
 SQ SEQUENCE 3898 AA; 438268 MW; D167BF5E48B11747 CRC64;

Query Match 50.0%; Score 49; DB 14; Length 3898;
 Best Local Similarity 55.6%; Pred. No. 51;
 Matches 10; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

OY 1 HKKNG---GKNGKNGW 14
 DB 3262 HKKNTLCPSGSLGKNGW 3279

RESULT 3
 ID 014262 PRELIMINARY; PRT; 225 AA.
 AC 014262;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-MAY-1999 (TREMblrel. 10, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE HYPOTHETICAL 26.6 KDA PROTEIN C7D4.05 IN CHROMOSOME I.
 GN SPAC7D4.05
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 AC Gentles S., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
 RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO YEAST YMR130W.
 DE EMBL: Z99532; CAB16722.2; -;
 DR InterPro: IPR001454; -;
 DR Pfam: PF00702; Hydrolase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 225 AA; 26299 MW; B6C38F28D230A2ED CRC64;

Query Match 48.0%; Score 47; DB 3; Length 225;

Best Local Similarity 56.2%; Pred. No. 6.2;
 Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

OY 1 HKKNGKNG--KNGW 14
 DB 60 HKKNGKSGLNPDMW 75

RESULT 4
 ID 09S207 PRELIMINARY; PRT; 275 AA.
 AC 09S207;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, last annotation update)
 DE PUTATIVE INTEGRAL MEMBRANE PROTEIN.
 GN SC466.04c.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed-8843436;
 RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL096884; CAB51427.1; -;
 SQ SEQUENCE 275 AA; 29424 MW; B46AF89DCA186591 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 275;
 Best Local Similarity 87.5%; Pred. No. 7.6;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 KNGKNGW 14
 DB 222 KNGKNGSW 229

RESULT 5
 ID 09HL50 PRELIMINARY; PRT; 95 AA.
 AC 09HL50;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE HYPOTHETICAL PROTEIN TAO381.
 GN TAO381.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
 OC Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed-11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Fishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma

RT acidophilum.";
 RL Nature 407:508-513(2000).
 DR EMBL: AL445064; CAC11525.1;
 KW Hypothetical protein.
 SQ SEQUENCE 95 AA; 11025 MW; 10B5282A55FB868 CRC64;

Query Match 46.9%; Score 46; DB 1; Length 95;
 Best Local Similarity 60.0%; Pred. No. 3.8;
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 HKNGKNGKNGM 15
 DB 72 HKOLSKNGSKNGRK 86

RESULT 6
 O9P6E6 PRELIMINARY; PRT; 516 AA.

ID O9P6E6;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE RAD57 RELATED PROTEIN.
 GN B17C10.30.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;

RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
 RA Nykatura G., Mewes H.W., Mannhaupt G.,
 RL Submitted (May-2000) to the EMBL/genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (May-2000) to the EMBL/genbank/DBJ databases.
 DR EMBL: AL355926; CAB91223.1;
 DR InterPro: IPR001553;
 DR SMART: IPR003593;
 DR SMART: SMO0382; AAA; 1.
 SQ SEQUENCE 516 AA; 55456 MW; 8A1B9CA95D6AA295 CRC64;

Query Match 46.9%; Score 46; DB 3; Length 516;
 Best Local Similarity 64.3%; Pred. No. 20;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 KNGKNGKNGM 15
 DB 486 KGVGKNGKNGM 499

RESULT 7
 O34507 PRELIMINARY; PRT; 648 AA.

AC O34507;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE YLOP PROTEIN.
 GN YLOP.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bartero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borriess R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.C., Comerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Enlian K.D., Erington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollighly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kletter-Blanchard M., Klein C.,
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Puje P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 RT subtilis."
 RT Nature 390:249-256(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,
 RL Submitted (Nov-1997) to the EMBL/genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Poulger D., Errington J.,
 RL Submitted (Aug-1997) to the EMBL/genbank/DBJ databases.
 CC -I- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: 299112; CAB13450.1;
 DR EMBL: Y13937; CAA74267.1;
 DR InterPro: IPR000719;
 DR InterPro: IPR002290;
 DR Pfam: PF00069; PKinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR SMART: SMO0220; S.TRC; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 648 AA; 71866 MW; 9653AB5CFBAA7900 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 648;
 Best Local Similarity 61.5%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 KNGKNGKNGM 14
 DB 317 ENKTKNGKNGM 329

RESULT 8
 O9SUVO PRELIMINARY; PRT; 857 AA.

AC O9SUVO;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE HYPOTHETICAL 96.9 KDA PROTEIN.
 GN F8B4.120 OR AT4G32420.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 RX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Terry N., Ardiles W., Buysshaert C., Dasseville R.,
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
 RA Villarroel R., Gielens J., Van Montagu M., Hohenseil J., Mewes H.W.,
 RA Mayer K.F.X., Lemcke K., Schueller C.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Terry N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,
 RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villarroel R.,
 RA Gielens J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL034567; CA22569.1; -;
 DR EMBL: AL161581; CAB79959.1; -;
 DR HSSP: P05092; 1CWL.
 DR InterPro: IPR001046; -;
 DR InterPro: IPR002130; -;
 DR Pfam: PF00160; pro_isomerase; 2.
 DR PRINTS: PR00153; CSAPISMRASE.
 DR PRODOM: PD001861; -; 1.
 DR PROSITE: PS0072; CSA_PPIASE_2; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 857 AA; 96854 MW; 17E4FA842D58D972 CRC64;

Query Match 46.9%; Score 46; DB 10; Length 857;
 Best Local Similarity 80.0%; Pred. No. 34;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KNGKKNKNGK 11
 ID 1: |||||
 DB 168 KSDGKNGK 177

RESULT 9
 ID 09HAW1 PRELIMINARY; PRT; 1372 AA.
 AC 09HAW1;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE III TRANSCRIPTION INITIATION FACTOR B' SHORT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20496900; PubMed-11040218;
 RA Schramm L., Pendergrast P.S., Sun Y., Hernandez N.;
 RT "Different human TFIIB activities direct RNA polymerase III
 RT transcription from TATA-containing and TATA-less promoters.";
 RL Genes Dev. 14:2650-2663(2000).
 DR EMBL: AF298152; AAG30221.1; -;
 DR Initiation factor.
 SQ SEQUENCE 1372 AA; 154605 MW; C7416FD179610F22 CRC64;

Query Match 46.9%; Score 46; DB 4; Length 1372;
 Best Local Similarity 57.1%; Pred. No. 54;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 KNGKKNKNGK 15
 ID 1: |||||
 DB 338 KNGKKNKNGK 351

RESULT 10
 ID 09HAW2 PRELIMINARY; PRT; 1388 AA.
 AC 09HAW2;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE III TRANSCRIPTION INITIATION FACTOR B' .
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20496900; PubMed-11040218;
 RA Schramm L., Pendergrast P.S., Sun Y., Hernandez N.;
 RT "Different human TFIIB activities direct RNA polymerase III
 RT transcription from TATA-containing and TATA-less promoters.";
 RL Genes Dev. 14:2650-2663(2000).
 DR EMBL: AF298151; AAG30220.1; -;
 DR Initiation factor.
 SQ SEQUENCE 1388 AA; 156302 MW; C1A6011AB937D12F CRC64;

Query Match 46.9%; Score 46; DB 4; Length 1388;
 Best Local Similarity 57.1%; Pred. No. 54;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 KNGKKNKNGK 15
 ID 1: |||||
 DB 338 KNGKKNKNGK 351

RESULT 11
 ID 09H197 PRELIMINARY; PRT; 2187 AA.
 AC 09H197;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE TFNR PROTEIN (FRAGMENT).
 GN TFNR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Keller A.R., Herchenbach J., Raschke H., Wirth B.;
 RT "Transcription factor-like nuclear regulator (TFNR) is a large protein
 RT with 9 repeats of a novel 55-amino acid motif closely localized to the
 RT survival motor neuron gene.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ279120; CAC21448.1; JOINED.
 DR EMBL: AJ279121; CAC21448.1; JOINED.
 DR EMBL: AJ279122; CAC21448.1; JOINED.
 DR EMBL: AJ279123; CAC21448.1; JOINED.
 DR EMBL: AJ279124; CAC21448.1; JOINED.
 DR EMBL: AJ279125; CAC21448.1; JOINED.
 DR EMBL: AJ279126; CAC21448.1; JOINED.
 DR EMBL: AJ279127; CAC21448.1; JOINED.
 DR EMBL: AJ279128; CAC21448.1; JOINED.
 DR EMBL: AJ279129; CAC21448.1; JOINED.
 DR EMBL: AJ279130; CAC21448.1; JOINED.
 DR EMBL: AJ279131; CAC21448.1; JOINED.
 DR EMBL: AJ279132; CAC21448.1; JOINED.
 DR EMBL: AJ279133; CAC21448.1; JOINED.

DR EMBL: AJ279134; CAC21448.1; JOINED.
 DR EMBL: AJ279135; CAC21448.1; JOINED.
 DR EMBL: AJ279136; CAC21448.1; JOINED.
 DR EMBL: AJ279137; CAC21448.1; JOINED.
 DR EMBL: AJ279138; CAC21448.1; JOINED.
 DR EMBL: AJ279139; CAC21448.1; JOINED.
 DR EMBL: AJ279140; CAC21448.1; JOINED.
 DR EMBL: AJ279141; CAC21448.1; JOINED.
 DR EMBL: AJ279142; CAC21448.1; JOINED.
 DR EMBL: AJ279143; CAC21448.1; JOINED.
 DR EMBL: AJ279144; CAC21448.1; JOINED.
 DR EMBL: AJ279145; CAC21448.1; JOINED.
 DR EMBL: AJ279146; CAC21448.1; JOINED.
 DR EMBL: AJ279147; CAC21448.1; JOINED.
 DR EMBL: AJ279148; CAC21448.1; JOINED.
 DR EMBL: AJ279149; CAC21448.1; JOINED.
 DR EMBL: AJ279150; CAC21448.1; JOINED.
 FT NON_TER 2187
 SQ SEQUENCE 2187 AA; 245726 MW; 2FB083F5F82AFB55 CRC64;

Query Match 46.9%; Score 46; DB 4; Length 2187;
 Best Local Similarity 57.1%; Pred. No. 85;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 KNGKRGKNGKNGK 15
 DB 338 KNGKRGKNGKNGK 351

RESULT 12
 ID 09HCY0 PRELIMINARY; PRT: 2254 AA.

DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PUTATIVE TRANSCRIPTION FACTOR-LIKE NUCLEAR REGULATOR.
 GN TFNR.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP Kelter A.R., Herchenbach J., Wirth B.:
 RA "The transcription factor like nuclear regulator (TFNR) contains a
 RT novel 55-amino acid motif repeated 9 times and maps closely to SMN1."
 RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ238520; CAC04245.1; -
 SQ SEQUENCE 2254 AA; 252844 MW; F350E96F53F04CPE CRC64;

Query Match 46.9%; Score 46; DB 4; Length 2254;
 Best Local Similarity 57.1%; Pred. No. 88;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 KNGKRGKNGKNGK 15
 DB 338 KNGKRGKNGKNGK 351

RESULT 13
 ID 016637 PRELIMINARY; PRT: 1251 AA.

AC 016637;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE W09G10.4 PROTEIN.
 GN W09G10.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RT Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Bentley D., Goela D., Holmes A.;
 RT "The sequence of C. elegans cosmid W09G10."
 RT Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF016671; AAB66112.2; -
 DR InterPro: IPR000408; -
 DR InterPro: IPR002553; -
 DR Pfam: PF01602; Adaptin N; 1.
 DR PROSITE: PS00626; RCC1.2; UNKNOWN.1.
 SQ SEQUENCE 1251 AA; 133949 MW; C3B0B18DAEFLA38 CRC64;

Query Match 45.9%; Score 45; DB 5; Length 1251;
 Best Local Similarity 64.3%; Pred. No. 70;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 KNGKRGKNGKNGK 15
 DB 787 KNGKRGKNGKNGK 800

RESULT 14

ID 09K6X3 PRELIMINARY; PRT: 298 AA.

AC 09K6X3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE CELL-DIVISION PROTEIN.
 GN FTSX.

OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RA Takami H., Nakasone K., Takaki Y.;
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP001519; BAB07320.1; -
 SQ SEQUENCE 298 AA; 33374 MW; 5F86FB72BC1B62BB CRC64;

Query Match 44.9%; Score 44; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 HKNGKNGKNGKNGK 16
 DB 9 HVRGKNGKNGKNGK 24

RESULT 15
 ID 09FIM2 PRELIMINARY; PRT; 806 AA.
 AC 09FIM2;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE CELL DIVISION PROTEIN FTSH.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=99156233; PubMed=10048488;
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned pl and TAC clones.";
 RL DNA Res. 5:379-391(1998).
 DR EMBL:AB016885; BAB09632.1; -
 SQ SEQUENCE 806 AA; 87837 MW; 074470A8A1E6FAF9 CRC64;

Query Match 44.9%; Score 44; DB 10; Length 806;
 Best Local Similarity 50.0%; Pred. No. 65;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 HKNKGKNGKNGM 14
 :|:|:|:|:|:
 Db 108 NKDKGRGKNGKELM 121

Search completed: July 6, 2001, 09:25:57
 Job time: 993 sec

GenCore version 4.5
(c) 1993 - 2000 Comp

el

Search time 113.68 Seconds
(without alignments)
6.399 Million cell updates/sec

PT -
XX
XX Claim 5: Page 27-28; 52pp; English.
XX
XX The present sequence is derived from human high molecular
CC weight kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds
CC with high affinity to endothelial cells, where it is cleaved to
CC two-chain high molecular weight kininogen (Hka) by plasma kallikrein.
CC Hka or a synthetic compound comprising part or all of the present
CC sequence may be used in a pharmaceutical composition for inhibiting
CC angiogenesis. Angiogenesis occurs in a number of disease states, such
CC as tumour formation and expansion, and certain ocular disorders. It can
CC also occur in a rheumatoid joint, hastening joint destruction by
CC allowing an influx of leukocytes. The composition may inhibit
CC angiogenesis by inhibiting endothelial cell proliferation or by
CC inducing endothelial cell apoptosis. Peptides used in the composition
CC may be recombinant peptides, natural peptides, or synthetic peptides.
CC They may also be chemically synthesised, using, for example, solid
CC phase synthesis methods.
XX
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 75; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGLGHC 12
| | | | | | | | | |
Db 1 hgneqghyghg 12

RESULT 2

AAW07627
ID AAW07627 standard; peptide; 17 AA.

XX
XX AAW07627;

DF 04-FEB-1997 (first entry)

DE Human high polymer quininogen L-chain derived peptide.

KW Human; high polymer; quininogen; L-chain.

OS Homo sapiens.

PN JP08208692-A

PD 13-AUG-1996.

PF 28-SEP-1995; 95JP-0276418.

PR 28-SEP-1994; 94JP-0259451.

PA (SUMU) SUMITOMO SEIYAKU KK.

DR WPI; 1996-421988/42.

XX Cell adhesion inhibiting peptide(s), used as cancer metastasis
PT inhibitor - comprises partial amino acid sequence of human high
PT polymer quininogen L chain
XX

PS Example; Page 8; 14pp; Japanese.

CC The present peptide is derived from residues 402-498 of the human
CC high polymer quininogen L-chain. It was synthesised using a solid
CC phase method, and purified using a YMC-DOS-120A-S15/13 column.
XX

SQ Sequence 17 AA;

Query Match 100.0%; Score 75; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGLGHC 12
| | | | | | | | | |
Db 4 hgneqghyghg 15

RESULT 3

AAW81996
ID AAW81996 standard; peptide; 28 AA.

XX
XX AAW81996;

DT 16-OCT-2000 (first entry)

DE Human high molecular weight kininogen domain 5 fragment #5.

KW Human; high molecular weight kininogen; HK;

KW two-chain high molecular weight kininogen; Hka;

KW angiogenesis inhibition; tumour; cancer; ocular disorder;

XX Rheumatoid arthritis; endothelial cell apoptosis.

OS Homo sapiens.

PN WO200027866-A1.

PD 18-MAY-2000.

PF 05-NOV-1999; 99WO-US26419.

PR 10-NOV-1998; 98US-0107833.

PA (UTEM) UNIV TEMPLE.

PI (MCCR/) MCCRAE R K.

DR WPI; 2000-376483/32.

XX A pharmaceutical composition used to inhibit angiogenesis, inhibit
PT endothelial cell proliferation, and induce endothelial cell apoptosis
PT
XX
XX Claim 8; Page 28; 52pp; English.

CC The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain
CC high molecular weight kininogen (Hka) by plasma kallikrein. Hka or a
CC synthetic compound comprising the present sequence may be used in a
CC pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
CC occurs in a number of disease states, such as tumour formation and
CC expansion, and certain ocular disorders. It can also occur in a
CC rheumatoid joint, hastening joint destruction by allowing
CC an influx of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial
CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.
XX

SQ Sequence 28 AA;

Query Match 100.0%; Score 75; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGLGHC 12
| | | | | | | | | |
Db 1 hgneqghyghg 12


```

RESULT 4
AAV93346
ID AAV93346 standard; peptide; 55 AA.
XX
XX
AC AAV93346;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
XX plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
XX endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
XX Homo sapiens.
XX
PN MO200027415-A2.
XX
PD 18-MAY-2000.
XX
PE 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM ) UNIT TEMPLE.
PA (DUPO ) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOSA/) MOUSA A S.
PI Colman WR, Mousa AS;
DR WPI: 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX
XX Claim 4; Page 36; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120-kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
XX Sequence 55 AA:
XX
OY 1 HGEEQHGSLGNG 12
Db 26 hgheqngslgng 37
XX
XX Query Match 100.0%; Score 75; DB 21; Length 55;
XX Best Local Similarity 100.0%; Pred. No. 7e-05;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
RESULT 5
AAV93348
ID AAV93348 standard; peptide; 62 AA.
XX
XX AAV93348;
XX
AC AAV93348;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.

```

XX	Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW	plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KM	endothelial cell proliferation; endothelial cell migration; vitronectin.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
PN	WO200027415-A2.
PD	18-MAY-2000.
XX	
PF	09-NOV-1999; 99WO-US26377.
XX	
PR	10-NOV-1998; 98US-0107844.
XX	
PA	(UTEM) UNIV TEMPLE.
PA	(DUPO) DUPONT PHARM CO.
PA	(COLM/) COLMAN W R.
PA	(MOOS/) MOUSA A S.
PI	Colman WR, Mousa AS;
XX	
DR	WPI; 2000-376306/32.
XX	
PT	Method for inhibiting endothelial cell proliferation, using compound
PT	that inhibit endothelial cell migration
XX	
PS	Claim 6; Page 37; 41pp; English.
XX	
CC	The present sequence represents an analogue of the light chain of human
CC	high molecular weight kininogen. High molecular weight kininogen is a
CC	120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC	where it is cleaved by plasma kallikrein into heavy and light chains.
CC	Analogues of high molecular weight kininogen are used in the method
CC	of the invention. The specification describes a method of inhibiting
CC	endothelial cell proliferation. The method comprises contacting
CC	endothelial cells with a compound containing high molecular weight
CC	kininogen analogues. The method and the compounds can be used for
CC	inhibiting endothelial cell proliferation. The compounds can also be
CC	used for inhibiting angiogenesis. The compounds can also be used to
CC	inhibit migration of endothelial cells to vitronectin.
XX	
SO	Sequence 62 AA;
XX	
Query Match	100.0%; Score 75; DB 21; Length 62;
Best Local Similarity	100.0%; Pred. No. 7, 9e-05;
Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 HGHEQOGLGHC 12
DB	5 hghgqghlgbyb 16
XX	
RESULT 6	
ID	AAR75186
XX	AAR75186 standard; peptide; 63 AA.
XX	AAR75186;
XX	05-DEC-1995 (first entry)
DE	Partial peptide of human HMW kininogen fragment 2.
XX	
KW	high molecular weight; kininogen; fragment; 1,2; 1; 2; partial;
KM	wound treating agent; bovine; growth promotion; fibroblast.
XX	
OS	Homo sapiens.
XX	
PN	JP07082172-A.
XX	
DD	28-MAR-1995.

XX 17-SEP-1993; 93JP-0230616.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX (FARH) HOECHST JAPAN KK.
XX
XX WPI; 1995-158909/21.
XX
XX A wound treating agent contg. a partial peptide of kininogen -
XX have growth promotion activity of fibroblasts.
XX
XX Claim 8; Page 8; 8pp; Japanese.
XX
XX AAR75186 is a partial peptide corresponding to human kininogen
XX fragment 1, amino acids 458-520. Partial peptides of bovine and
XX human kininogen fragments 1,2, 1 and 2, are used in wound treating
XX agent compns. and act as the active component. The fragments are
XX useful in wound treating because they have growth promotion activity
XX on fibroblasts.
XX
XX Sequence 63 AA;
XX
XX

Query Match 100.0%; Score 75; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGGLGSG 12
| | | | | | | | | | | | | |
Db 6 hgheqghg1ghg 17

RESULT 7

AA93347
ID AAY93347 standard; peptide; 83 AA.

XX
AC AAY93347;

XX
DT 04-SEP-2000 (first entry)

XX
DE Light chain of human high molecular weight kininogen analogue.

XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiotensin;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX
OS Synthetic.

XX
OS Homo sapiens.

XX
PN WO200027415-A2.

XX
PD 18-MAY-2000.

XX
PE 09-NOV-1999; 99WO-US26377.

XX
PR 10-NOV-1998; 98US-0107844.

XX
PA (UTEM) UNIV TEMPLE.

XX
PA (DUPO) DUPONT PHARM CO.

XX
PA (COLM/) COLMAN W R.

XX
PA (MOUS/) MOUSA A S.

XX
PI Colman WR, Mousa AS;

XX
DR WPI; 2000-376306/32.

XX
PT Method for inhibiting endothelial cell proliferation, using compound
XX that inhibit endothelial cell migration -

XX
PS Claim 5; Page 37; 41pp; English.

XX
CC The present sequence represents an analogue of the light chain of human

CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiotensin. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.

XX
XX Sequence 83 AA;
XX
XX

Query Match 100.0%; Score 75; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGGLGSG 12
| | | | | | | | | | | | | |
Db 26 hgheqghg1ghg 37

RESULT 8
AA93351
ID AAY93351 standard; peptide; 94 AA.

XX
AC AAY93351;

XX
DT 04-SEP-2000 (first entry)

XX
DE Light chain of human high molecular weight kininogen analogue.

XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiotensin;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX
OS Synthetic.

XX
OS Homo sapiens.

XX
PN WO200027415-A2.

XX
PD 18-MAY-2000.

XX
PE 09-NOV-1999; 99WO-US26377.

XX
PR 10-NOV-1998; 98US-0107844.

XX
PA (UTEM) UNIV TEMPLE.

XX
PA (DUPO) DUPONT PHARM CO.

XX
PA (COLM/) COLMAN W R.

XX
PA (MOUS/) MOUSA A S.

XX
PI Colman WR, Mousa AS;

XX
DR WPI; 2000-376306/32.

XX
PT Method for inhibiting endothelial cell proliferation, using compound
XX that inhibit endothelial cell migration -
XX
XX Claim 8; Page 39; 41pp; English.

XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be

CC used for inhibiting angiogenesis. The compounds can also be used to
inhibit migration of endothelial cells to vitronectin.

SO Sequence 94 AA:

Query Match 100.0%; Score 75; DB 21; Length 94;

Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGHEQOHGIGHG 12
|||||
DB 26 hgheqghg1ghg 37

RESULT 9

AA75181
ID AAR75181 standard; peptide; 131 AA.

AC AAR75181;

DT 05-DEC-1995 (first entry)

DE Partial peptide of human HMW kininogen fragment 1.2.

XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;

KW wound treating agent; human; growth promotion; fibroblast.

OS Homo sapiens.

PN JP07082172-A.

PD 28-MAR-1995.

PF 17-SEP-1993; 93JP-0230616.

PR 17-SEP-1993; 93JP-0230616.

PA (FARH) HOECHST JAPAN KK.

DR WPI: 1995-158909/21.

PT A wound treating agent contg. a partial peptide of kininogen -
have growth promotion activity of fibroblasts.

PS Claim 7; Page 7; 8pp; Japanese.

CC AAR75181 is a partial peptide corresponding to human kininogen
fragment 1.2, amino acids 390-520. Partial peptides of bovine and
human kininogen fragments 1.2, 1 and 2, are used in wound treating
agent compsns. and act as the active component. The fragments are
useful in wound treating because they have growth promotion activity
on fibroblasts.

SO Sequence 131 AA:

Query Match 100.0%; Score 75; DB 16; Length 131;

Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGHEQOHGIGHG 12
|||||
DB 74 hgheqghg1ghg 85

RESULT 10

AA75181
ID AAY93349 standard; peptide; 186 AA.

AC AAY93349;

DT 04-SEP-2000 (first entry)

XX Light chain of human high molecular weight kininogen analogue.

DE Human; high molecular weight kininogen; glycoprotein; endothelial cell;

KW Plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

KW endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.

OS Homo sapiens.

PN WO200027415-A2.

PD 18-MAY-2000.

PF 09-NOV-1999; 99WO-US26377.

PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOSA/) MOUSA A S.

PI Colman WR, Mousa AS;

DR WPI: 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound
that inhibit endothelial cell migration

PS Claim 9; Page 38; 41pp; English.

CC The present sequence represents an analogue of the light chain of human
high molecular weight kininogen. High molecular weight kininogen is a
120 kDa glycoprotein which binds with high affinity to endothelial cells,
where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
of the invention. The specification describes a method of inhibiting
endothelial cell proliferation. The method comprises contacting
endothelial cells with a compound containing high molecular weight
kininogen analogues. The method and the compounds can be used for
inhibiting endothelial cell proliferation. The compounds can also be
used for inhibiting angiogenesis. The compounds can also be used to
inhibit migration of endothelial cells to vitronectin.

SO Sequence 186 AA:

Query Match 100.0%; Score 75; DB 21; Length 186;

Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGHEQOHGIGHG 12
|||||
DB 5 hgheqghg1ghg 16

RESULT 11

AA75181
ID AAY93342 standard; protein; 255 AA.

AC AAY93342;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;

KW Plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

KW endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Homo sapiens.

OS

PN WO200027415-A2.
 XX 18-MAY-2000.
 XX 09-NOV-1999; 99WO-US26377.
 XX 10-NOV-1998; 98US-0107844.
 PR (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX Colman WR, Mousa AS;
 PI WPI: 2000-376306/32.
 DR Method for inhibiting endothelial cell proliferation, using compound
 PT that inhibit endothelial cell migration
 XX
 XX Disclosure; Page 3; 41pp; English.
 PS
 CC The present sequence represents the light chain of human high molecular
 CC weight kininogen. High molecular weight kininogen is a 120 kDa
 CC glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method
 CC of the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be
 CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.
 XX
 SQ Sequence 255 AA;

Query Match. 100.0%; Score 75; DB 21; Length 255;
 Best Local Similarity 100.0%; Pred. No. 0.00033;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGIGHG 12
 |||||||||
 Db 74 hgneqghlgh 85

RESULT 12
 AAY93350
 ID AAY93350 standard; peptide: 16 AA.
 XX
 XX AAY93350;
 AC
 XX
 XX 04-SEP-2000 (first entry)
 DT
 XX
 DE Light chain of human high molecular weight kininogen analogue.
 XX
 KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200027415-A2.
 XX 18-MAY-2000.
 PD
 XX 09-NOV-1999; 99WO-US26377.
 PF
 XX 10-NOV-1998; 98US-0107844.
 PR
 XX (UTEM) UNIV TEMPLE.
 PA

PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX Colman WR, Mousa AS;
 PI WPI: 2000-376306/32.
 DR Method for inhibiting endothelial cell proliferation, using compound
 PT that inhibit endothelial cell migration
 XX
 XX Claim 7; Page 39; 41pp; English.
 PS
 CC The present sequence represents an analogue of the light chain of human
 CC high molecular weight kininogen. High molecular weight kininogen is a
 CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
 CC where it is cleaved by plasma kallikrein into heavy and light chains.
 CC Analogues of high molecular weight kininogen are used in the method
 CC of the invention. The specification describes a method of inhibiting
 CC endothelial cell proliferation. The method comprises contacting
 CC endothelial cells with a compound containing high molecular weight
 CC kininogen analogues. The method and the compounds can be used for
 CC inhibiting endothelial cell proliferation. The compounds can also be
 CC used for inhibiting angiogenesis. The compounds can also be used to
 CC inhibit migration of endothelial cells to vitronectin.
 XX
 SQ Sequence 16 AA;

Query Match. 92.0%; Score 69; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.00016;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGIGHG 11
 |||||||||
 Db 6 hgneqghlgh 16

RESULT 13
 AAY93343
 ID AAY93343 standard; peptide: 186 AA.
 XX
 XX AAY93343;
 AC
 XX
 XX 04-SEP-2000 (first entry)
 DT
 XX
 DE Light chain of human high molecular weight kininogen analogue.
 XX
 KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
 KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
 KW endothelial cell proliferation; endothelial cell migration; vitronectin.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Misc-difference 1..186
 FT /note="Xaa are unspecified amino acids"
 XX
 PN WO200027415-A2.
 XX 18-MAY-2000.
 PD
 XX 09-NOV-1999; 99WO-US26377.
 PF
 XX 10-NOV-1998; 98US-0107844.
 PR
 XX (UTEM) UNIV TEMPLE.
 PA (DUPO) DUPONT PHARM CO.
 PA (COLM/) COLMAN W R.
 PA (MOUS/) MOUSA A S.
 XX Colman WR, Mousa AS;
 PI

```

XX DR WPI: 2000-376306/32.
XX
XX PT Method for inhibiting endothelial cell proliferation, using compound
XX PT that inhibit endothelial cell migration
XX PS Claim 1: Page 34-35; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 186 AA:

Query Match          92.0%; Score 69; DB 21; Length 186;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGLGH 11
   |||||
DB 171 hghqgqghgh 181

RESULT 14
AAR75180
ID AAR75180 standard; peptide; 41 AA.
XX
AC AAR75180;
XX
DT 05-DEC-1995 (first entry)
XX
DE Partial peptide of HMW kininogen fragment 2.
XX
KM high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KM wound treating agent; bovine; growth promotion; fibroblast.
XX
OS Bos taurus.
XX
PN JP07082172-A.
XX
PD 28-MAR-1995.
XX
PF 17-SEP-1993; 93JP-0230616.
XX
PR 17-SEP-1993; 93JP-0230616.
XX
PA (FARH ) HOECHST JAPAN KK.
XX
DR WPI: 1995-158909/21.
XX
PT A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
PS Claim 6; Page 7; 8pp; Japanese.
XX
CC AAR75179 is a partial peptide corresponding to bovine kininogen
CC fragment 2, amino acids 456-496. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent comps. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 41 AA:

```

```

Query Match          90.7%; Score 68; DB 16; Length 41;
Best Local Similarity 83.3%; Pred. No. 0.00059;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGLGH 12
   |||||
DB 6 hghqgqghgh 17

RESULT 15
AAR75178
ID AAR75178 standard; peptide; 110 AA.
XX
AC AAR75178;
XX
DT 05-DEC-1995 (first entry)
XX
DE Partial peptide of HMW kininogen fragment 1.2.
XX
KM high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KM wound treating agent; bovine; growth promotion; fibroblast.
XX
OS Bos taurus.
XX
PN JP07082172-A.
XX
PD 28-MAR-1995.
XX
PF 17-SEP-1993; 93JP-0230616.
XX
PR 17-SEP-1993; 93JP-0230616.
XX
PA (FARH ) HOECHST JAPAN KK.
XX
DR WPI: 1995-158909/21.
XX
PT A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
PS Claim 4; Page 6; 8pp; Japanese.
XX
CC AAR75178 is a partial peptide corresponding to bovine kininogen
CC fragment 1.2, amino acids 387-496. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent comps. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
SQ Sequence 110 AA:

Query Match          90.7%; Score 68; DB 16; Length 110;
Best Local Similarity 83.3%; Pred. No. 0.0016;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEQOHGLGH 12
   |||||
DB 75 hghqgqghgh 86

```

Search completed: July 6, 2001, 09:09:16
Job time: 122 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:07:14 ; Search time 56.74 Seconds
(without alignments)
4.260 Million cell updates/sec

Title: US-09-437-912-1

Sequence: 1 HGEHQHGLGHC 12

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCUTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	53.3	154	1	US-08-450-945-67 Sequence 67, Appl
2	40	53.3	154	4	US-08-976-161-67 Sequence 67, Appl
3	39	52.0	417	1	US-08-553-703A-2 Sequence 2, Appl
4	39	52.0	417	2	US-09-006-021-2 Sequence 2, Appl
5	39	52.0	419	1	US-08-553-703A-3 Sequence 3, Appl
6	39	52.0	419	2	US-09-006-021-3 Sequence 3, Appl
7	39	52.0	420	2	US-08-845-998-8 Sequence 8, Appl
8	39	52.0	420	4	US-09-206-537-8 Sequence 8, Appl
9	38	50.7	113	2	US-08-918-727-7 Sequence 7, Appl
10	38	50.7	113	3	US-09-205-680A-7 Sequence 7, Appl
11	38	50.7	286	3	US-09-203-716-1 Sequence 1, Appl
12	37.5	50.0	60	1	US-08-255-457-1 Sequence 1, Appl
13	37.5	50.0	60	2	US-09-115-032-1 Sequence 1, Appl
14	37.5	50.0	60	5	PCT-US95-05772-1 Sequence 1, Appl
15	36.5	48.7	1958	1	US-07-945-283-2 Sequence 2, Appl
16	36	48.0	43	2	US-08-488-161-73 Sequence 73, Appl
17	36	48.0	43	3	US-09-273-685-73 Sequence 73, Appl
18	36	48.0	43	5	PCT-US95-11934-73 Sequence 73, Appl
19	36	48.0	131	1	US-07-779-706A-2 Sequence 2, Appl
20	36	48.0	131	1	US-08-104-503-2 Sequence 2, Appl
21	36	48.0	708	3	US-08-965-903B-8 Sequence 8, Appl
22	36	48.0	708	1	US-08-396-479B-8 Sequence 8, Appl
23	36	48.0	708	1	US-08-818-823-8 Sequence 8, Appl
24	36	48.0	739	1	US-08-396-479B-10 Sequence 10, Appl
25	36	48.0	739	1	US-08-818-823-10 Sequence 10, Appl
26	36	48.0	1068	1	US-08-396-479B-12 Sequence 12, Appl
27	36	48.0	1068	1	US-08-818-823-12 Sequence 12, Appl

28	36	48.0	1075	5	PCT-US94-07297-41	Sequence 41, Appl
29	36	48.0	1250	1	US-08-441-139-9	Sequence 9, Appl
30	35	46.7	64	1	US-08-200-016-6	Sequence 6, Appl
31	35	46.7	109	1	US-07-987-272A-8	Sequence 8, Appl
32	35	46.7	114	1	US-08-385-241-3	Sequence 3, Appl
33	35	46.7	347	1	US-07-637-870-5	Sequence 5, Appl
34	35	46.7	347	1	US-07-640-476-10	Sequence 10, Appl
35	35	46.7	348	1	US-07-637-399-8	Sequence 8, Appl
36	35	46.7	348	1	US-08-112-703-8	Sequence 8, Appl
37	35	46.7	387	1	US-07-637-870-3	Sequence 3, Appl
38	35	46.7	387	1	US-07-637-399-9	Sequence 9, Appl
39	35	46.7	387	1	US-07-640-476-12	Sequence 12, Appl
40	35	46.7	387	1	US-08-112-703-9	Sequence 9, Appl
41	35	46.7	388	1	US-07-637-870-4	Sequence 4, Appl
42	35	46.7	388	1	US-07-637-399-7	Sequence 7, Appl
43	35	46.7	388	1	US-07-640-476-7	Sequence 7, Appl
44	35	46.7	388	1	US-08-112-703-7	Sequence 7, Appl
45	35	46.7	389	1	US-07-640-476-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-450-945-67
Sequence 67, Application US/08450945
Patent No. 5783383
GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhiko
APPLICANT: MocarSKI, Edward S. Jr.
TITLE OF INVENTION: LATENT TRANSCRIPTS AND PROMOTERS
TITLE OF INVENTION: OF CYTOMEGALOVIRUS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dellinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,945
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles R.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 154 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-450-945-67

Query Match 53.3%, Score 40; DB 1; Length 154;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGEHQHGLGHC 12
DB 61 HSLQRRGLGHC 72

RESULT 2
US-08-976-161-67
Sequence 67, Application US/08976161
Patent No. 6194542
GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhiro
TITLE OF INVENTION: LATENT TRANSCRIPTS AND PROMOTERS
TITLE OF INVENTION: OF CYTOMEGALOVIRUS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,161
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/450,945
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 154 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-161-67

Query Match 53.3%; Score 40; DB 4; Length 154;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 HGHQDQHG 12
DB 61 HSLQERRGLGHG 72

RESULT 3
US-08-553-703A-2
Sequence 2, Application US/08553703A
Patent No. 5795767
GENERAL INFORMATION:
APPLICANT: MARU, ISAFUMI
APPLICANT: OHTA, YASUHIRO
APPLICANT: TSUKADA, YOJI
TITLE OF INVENTION: EPIMERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,703A
FILING DATE: 30-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-553-703A-2

Query Match 52.0%; Score 39; DB 1; Length 417;
Best Local Similarity 62.5%; Pred. No. 95;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGHQDQHG 8
DB 29 HSHDQEHG 36

RESULT 4
US-09-006-021-2
Sequence 2, Application US/09006021
Patent No. 5994105
GENERAL INFORMATION:
APPLICANT: MARU, ISAFUMI
APPLICANT: OHTA, YASUHIRO
APPLICANT: TSUKADA, YOJI
TITLE OF INVENTION: EPIMERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,021
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/553,703
FILING DATE: 30-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:

TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-09-006-021-2

Query Match 52.0%; Score 39; DB 2; Length 417;
Best Local Similarity 62.5%; Pred. No. 95;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 HGHEQOHG 8
| | | | |
Db 29 HSHDQEHG 36

RESULT 5
US-08-553-703A-3
; Sequence 3, Application US/08553703A
; Patent No. 5795767
; GENERAL INFORMATION:
; APPLICANT: MARU, ISAFUMI
; APPLICANT: OHTA, YASUHIRO
; APPLICANT: TSUKADA, YOJI
; TITLE OF INVENTION: EPIMERASE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,703A
; FILING DATE: 30-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-553-703A-3

Query Match 52.0%; Score 39; DB 1; Length 419;
Best Local Similarity 62.5%; Pred. No. 95;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Oy 1 HGHEQOHG 8
| | | | |
Db 29 HSHDQEHG 36

RESULT 6
US-09-006-021-3
; Sequence 3, Application US/09006021
; Patent No. 5994105
; GENERAL INFORMATION:
; APPLICANT: MARU, ISAFUMI
; APPLICANT: OHTA, YASUHIRO
; APPLICANT: TSUKADA, YOJI
; TITLE OF INVENTION: EPIMERASE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,021
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/553,703
; FILING DATE: 30-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-09-006-021-3

Query Match 52.0%; Score 39; DB 2; Length 419;
Best Local Similarity 62.5%; Pred. No. 95;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 HGHEQOHG 8
| | | | |
Db 29 HSHDQEHG 36

RESULT 7
US-08-845-998-8
; Sequence 8, Application US/08845998
; Patent No. 5879892
; GENERAL INFORMATION:
; APPLICANT: Van Baren, Nicolas
; APPLICANT: Coulle, Pierre G.
; APPLICANT: De Smet, Charles
; APPLICANT: Lucas, Sophie

APPLICANT: Boon, Thierry
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,998
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)720-3500
TELEFAX: (617)720-2441
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-845-998-8

Query Match 52.0%; Score 39; DB 2; Length 420;
Best Local Similarity 54.5%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HGHEQOGLGH 11
| | | | |
Db 189 HPHPHMSLGH 199

RESULT 8
US-09-206-537-8
Sequence 8, Application US/09206537
Patent No. 6130052
GENERAL INFORMATION:
APPLICANT: Van Baren, Nicolas
APPLICANT: Coulle, Pierre G.
APPLICANT: De Smet, Charles
APPLICANT: Lucas, Sophie
APPLICANT: Boon, Thierry
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/206,537
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/845,998
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)720-3500
TELEFAX: (617)720-2441
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-206-537-8

Query Match 52.0%; Score 39; DB 4; Length 420;
Best Local Similarity 54.5%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HGHEQOGLGH 11
| | | | |
Db 189 HPHPHMSLGH 199

RESULT 9
US-08-918-727-7
Sequence 7, Application US/08918727
Patent No. 5849528
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lai, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN S100 PROTEINS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,727
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: Pf-0373 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:

LIBRARY: GenBank
CLONE: 488157
US-08-918-727-7

Query Match 50.7%; Score 38; DB 2; Length 113;
Best Local Similarity 54.5%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 GHEQHGGLGHC 12
11:11111
DB 102 GHDRHGKCG 112

RESULT 10
US-09-205-680A-7
Sequence 7, Application US/09205680A
Patent No. 6103497
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN S100 PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/205,680A
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Colette C. Muenzen
REGISTRATION NUMBER: 39,784
REFERENCE/DOCKET NUMBER: PF-0373 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 488157
US-09-205-680A-7

Query Match 50.7%; Score 38; DB 3; Length 113;
Best Local Similarity 54.5%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 GHEQHGGLGHC 12
11:11111
DB 102 GHDRHGKCG 112

RESULT 11
US-09-203-716-1
Sequence 1, Application US/09203716
Patent No. 6001653
GENERAL INFORMATION:
APPLICANT: Crooke, Stanley T.
APPLICANT: Lima, Walter F.
APPLICANT: Wu, Hongjiang
TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
FILE REFERENCE: ISPH-0333
CURRENT APPLICATION NUMBER: US/09/203,716
CURRENT FILING DATE: 1998-12-02
EARLIER APPLICATION NUMBER: 60/067,458
EARLIER FILING DATE: 1997-12-04
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 286
TYPE: prt
ORGANISM: Homo sapiens
US-09-203-716-1

Query Match 50.7%; Score 38; DB 3; Length 286;
Best Local Similarity 85.7%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GHEQHG 8
111111
DB 82 GHEQHG 88

RESULT 12
US-08-235-457-1
Sequence 1, Application US/08255457
Patent No. 5780040
GENERAL INFORMATION:
APPLICANT: Plaut, Andrew G.
APPLICANT: Gilbert-Rotstein, Joanne V.
APPLICANT: Wright, Andrew
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/255,457
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-255-457-1

Query Match 50.0%; Score 37.5; DB 1; Length 60;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 3 HEOQHGLGH 11
11:1111 11
DB 4 HEOQH-GH 11

RESULT 13
US-09-115-032-1
Sequence 1, Application US/09115032
Patent No. 5972348
GENERAL INFORMATION:
APPLICANT: Plaut, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,032
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/255,457
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-115-032-1

Query Match 50.0%; Score 37.5; DB 2; Length 60;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 3 HEOQHGLGH 11
11:1111 11
DB 4 HEOQH-GH 11

RESULT 14
PCT-US95-05772-1
Sequence 1, Application PC/TUS9505772
GENERAL INFORMATION:

APPLICANT: Plaut, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.

TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL
TITLE OF INVENTION: BINDING PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05772
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-05772-1

Query Match 50.0%; Score 37.5; DB 5; Length 60;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 3 HEOQHGLGH 11
11:1111 11
DB 4 HEOQH-GH 11

RESULT 15
US-07-945-283-2
Sequence 2, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
TITLE OF INVENTION: Involving The EPO and LIT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911

CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Ribando, Curtis P
 REGISTRATION NUMBER: 27976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 309-685-4011 ext.513
 TELEFAX: 309-685-4128
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1958 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-945-283-2

Query Match 48.7%; Score 36.5; DB 1; Length 1958;
 Best Local Similarity 63.6%; Pred. No. 1e+03;
 Matches 7; Conservative 1; Mismatches 0; Indels 3; Gaps 1;
 QY 1 HGHEQOHGIGH 11
 || :||||
 Db 135 HG--EHGIGH 142

Search completed: July 6, 2001, 09:10:20
 Job time: 186 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:07:14 ; Search time 73.59 Seconds
(without alignments)
12.421 Million cell updates/sec

Title: US-09-437-912-1

Perfect score: 75

Sequence: 1 HGHEOQHGLGHC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	644	1 KGH0H1	kininogen, HMW pre
2	68	90.7	619	1 KGB0H1	kininogen, HMW II
3	68	90.7	621	1 KGB0H1	kininogen, HMW I P
4	62	82.7	264	2 C25486	K-kininogen, HMW P
5	55	73.3	639	2 A25486	kininogen, HMW I P
6	55	73.3	189	2 C81428	peptidyl-prolyl ci
7	55	73.3	290	2 C27115	K-kininogen, LMW P
8	54	72.0	315	2 A27115	major acute phase
9	54	72.0	314	2 T35241	hypothetical prote
10	52	69.3	85	2 A55969	hemolymph antifu
11	51	68.0	398	2 T02681	probable zinc tran
12	51	68.0	409	2 E83992	ATP/GTP-binding pr
13	50.5	67.3	110	2 T07618	cold stress protei
14	50	66.7	199	2 T48099	hypothetical prote
15	50	66.7	535	2 S66148	gene pipsqueak pro
16	50	66.7	1085	2 S66149	gene pipsqueak pro
17	49	65.3	457	2 S39079	puff C-8 protein -
18	49	65.3	670	2 F36791	hypothetical prote
19	49	65.3	735	2 T45059	hypothetical prote
20	48	65.3	2038	2 A43742	female sterile hom
21	48	64.0	439	2 S58327	cobalt accumulat
22	47	62.7	160	2 T07180	hypothetical prote
23	47	62.7	179	2 A85217	hypothetical prote
24	47	62.7	277	2 T04441	hypothetical prote
25	47	62.7	490	2 T36920	hypothetical prote
26	47	62.7	515	2 T23089	hypothetical prote
27	46.5	62.0	390	2 T34137	hypothetical prote
28	45	60.0	420	2 A49642	transcription fact
29	45	60.0	503	2 S54302	zinc transporter 2

30	45	60.0	507	2 S54303	zinc transport pro
31	45	60.0	529	2 T08684	hypothetical prote
32	44	58.7	208	2 T07732	tuberculosis-relat
33	44	58.7	356	2 H71496	probable aminopept
34	44	58.7	378	2 T49164	zinc transporter-1
35	44	58.7	389	2 B96635	hypothetical prote
36	44	58.7	690	2 H69268	copper-transportin
37	44	58.7	697	2 T03834	nuclear distributi
38	43.5	58.0	136	2 J02266	cold acclimation p
39	43.5	58.0	410	2 T26757	hypothetical prote
40	43	57.3	18	2 B32473	histidine-rich pro
41	43	57.3	203	2 T36240	hypothetical prote
42	43	57.3	335	2 D38532	hypB protein - Rho
43	43	57.3	507	2 D64575	hypothetical prote
44	42	56.0	102	2 T30119	hypothetical prote
45	42	56.0	177	2 S65780	glycine/proline-ri

ALIGNMENTS

RESULT 1
KGH0H1
kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-chiol proteinase inhibitor; preprokininogen; prokininogen
N:Contains: bradykinin (Kallidin I); HMW kininogen I; HMW kininogen II; low molecular
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030;
R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A:Title: Isolation of a human cDNA for alpha-2-chiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:85122621
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:K02566; NID:q177889
R:Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and 1
A:Reference number: A92344; MUID:85234582
A:Accession: A25276
A:Molecule type: mRNA
A:Residues: 1-592, 1, 594-644 <TAK>
A:Cross-references: GB:M1437; NID:g186751; PID:AA59550.1; PID:g386852
R:Auerwald, E.A.; Roessler, D.; Mentle, R.; Asfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A:Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:93223854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: ANSW, 253-377 <AUE>
A:Note: differences are due to known cloning artifacts
R:Loetspeich, F.; Kellermann, J.; Henschen, A.; Foerisch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kint
A:Reference number: A91153; MUID:86030270
A:Accession: A91153
A:Molecule type: Protein
A:Residues: 379-644 <LOT>
A:Note: the bradykinin sequence preceding the light chain sequence was not determined
R:Kellermann, J.; Loetspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A:Title: Completion of the primary structure of human high-molecular-mass kininogen.
A:Reference number: A24871; MUID:86108361
A:Accession: A24871
A:Molecule type: Protein
A:Residues: 27, 20-380 <KEU>
R:Kellermann, J.; Loetspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp 85-89, Plenum Press, New
A:Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A:Reference number: A27899
A:Accession: A27899

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEI2>
 R:Mindrou, T.; Carretero, O.A.; Prod' D.; Walz, D.; Scicch, A.G.
 A:Reference number: A91455; MUID:90255622
 A:Contents: annotation: bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
 C:Comment: The kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is a
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379, 390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F:19-131/Domain: cystatin homology <CY1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <BDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19-Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status expert
 F:28-614, 83-94, 107-126, 142-145, 206-218, 229-248, 264-267, 328-340, 351-370/Disulfide bond
 F:48/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:169, 205, 284/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:379-389/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 593, 628/Binding site: carbohydrate (Thr) (covalent) #status
 F:577/Binding site: carbohydrate (Ser) (covalent) #status experimental

A:Reference number: A61495; MUID:88211869
 A:Contents: annotation: bradykinin in human
 A:Molecule type: protein
 A:Residues: 1-19;189-192;310-314;381-389 <LENI>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869
 A:Contents: annotation: bradykinin in human
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55239; MUID:95251593
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Straczek, J.; Maschl, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; MUID:96033974
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evoluti
 A:Reference number: A92545; MUID:85234583
 A:Contents: annotation: gene organization

R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:90255622
 A:Contents: annotation: bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: The kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is a
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379, 390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F:19-131/Domain: cystatin homology <CY1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <BDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19-Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status expert
 F:28-614, 83-94, 107-126, 142-145, 206-218, 229-248, 264-267, 328-340, 351-370/Disulfide bond
 F:48/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:169, 205, 284/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:379-389/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 593, 628/Binding site: carbohydrate (Thr) (covalent) #status
 F:577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0% Score 75; DB 1; Length 644;
 Best Local Similarity 100.0% Pred. No. 0.00021;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGHEQDHGHLGHC 12
 Db 463 HGHEQDHGHLGHC 474

RESULT 2
 KGB0H2
 kininogen, HMW II precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence-revision 14-Nov-1983 #text-change 22-Jun-1999
 C:Accession: A01282; A91923; A91941; A91938; B29559
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A:Title: A single gene for bovine high molecular weight and low molecular weight kin
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01282
 A:Molecule type: mRNA
 A:Residues: 1-619 <Kin>
 A:Cross-references: GB:V01492; GB:K01758; NID:q493; PIDN:CAA24736.1; PID:q494
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 376-391 <KAT>
 R:Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.

J. Biochem. 79, 1201-1222, 1976
 A>Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino acid sequence has been determined from the complementary DNA sequence.
 A:Reference number: A91941; MUID:76260155
 A:Accession: A91941
 A:Molecule type: protein
 A:Residues: 387-455 <HAN>
 A:Note: 398-Pro, 401-Val, and 455-Lys were also found
 J. Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A>Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Amino acid sequence of the primary structure of bovine high-molecular-weight kininogen. Amino acid sequence of the primary structure of bovine high-molecular-weight kininogen.
 A:Reference number: A91938; MUID:75170265
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 456-496 <HAN>
 R. Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga, S.
 J. Biol. Chem. 262, 2768-2779, 1987
 A>Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of cleavage sites, and the primary structure of bovine high-molecular-weight kininogen.
 A:Reference number: A92627; MUID:87137530
 A:Accession: B29559
 A:Molecule type: protein
 A:Residues: 120-104, 106-256, 257-376 <SUE>
 R. Lotz, F.; Kellermann, J.; Henschen, A.; Foerisch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A>Title: The amino acid sequence of the light chain of human high-molecular-weight kininogen.
 A:Reference number: A91153; MUID:86030270
 A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R. Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 808, 1984
 A>Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation; disulfide bonds
 A:Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as the LMW form.
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is implicated in the release of bradykinin.
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F:1-18/Domain: signal sequence #status predicted <SUE>
 F:19-619/Product: HMW kininogen II #status predicted <MAT>
 F:19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
 F:19-130/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:261-372/Domain: cystatin homology <CY3>
 F:377-386/Product: lysyl-bradykinin (kallidin I) #status experimental <KBDV>
 F:378-386/Product: bradykinin (kallidin II) #status experimental <KBDV>
 F:387-619/Product: HMW kininogen II #status experimental <BDV>
 F:418-488/Region: glycine/histidine/lysine-rich
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
 F:27-589, 82-93, 106-125, 141-144, 205-217, 228-247, 261-264, 325-337, 348-367/Disulfide bonds:
 F:87/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:168, 169, 204, 280/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
 F:19/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:380/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:396, 400, 404, 510/Binding site: carbohydrate (Ser) (covalent) #status experimental
 F:397, 398, 518, 522, 534, 546, 551, 568/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F:496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 90.7% Score 68; DB 1; Length 619;
 Best Local Similarity 83.3%; Pred. No. 0.0026;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHEOQGLGHS 12
 |||:|||||
 DB 461 HGHOQGLGHS 472

RESULT 3

KGB0H1
 kininogen, HMW I precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A01281; A91923; A91938; A29559
 R. Kikamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakaiishi, S.
 Nature 305, 545-549, 1983
 A>Title: A single gene for bovine high molecular weight and low molecular weight kininogen.
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01281
 A:Molecule type: mRNA
 A:Residues: 1-621 <KIT>
 A:Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
 R. Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A>Title: Studies on the primary structure of bovine kininogen: cleavages of disulfide bonds a
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 378-393 <KAT>
 R. Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A>Title: Studies on the primary structure of bovine high-molecular-weight kininogen.
 A:Reference number: A91938; MUID:75170265
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 458-498 <HAN>
 R. Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga, S.
 J. Biol. Chem. 262, 2768-2779, 1987
 A>Title: Bovine high molecular weight kininogen. The amino acid sequence, positions o
 A:Reference number: A92627; MUID:87137530
 A:Accession: A29559
 A:Molecule type: protein
 A:Residues: 120-123, 125-127, 129-378 <SUE>
 R. Lotz, F.; Kellermann, J.; Henschen, A.; Foerisch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A>Title: The amino acid sequence of the light chain of human high-molecular-weight kininogen.
 A:Reference number: A91153; MUID:86030270
 A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R. Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 808, 1984
 A>Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation; disulfide bonds
 A:Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form
 C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is implicated in the release of bradykinin.
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status predicted <SUE>
 F:19-621/Product: HMW kininogen I #status predicted <MAT>
 F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
 F:19-130/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:263-374/Domain: cystatin homology <CY3>
 F:379-388/Product: lysyl-bradykinin (kallidin I) #status experimental <BDV>
 F:380-621/Product: bradykinin (kallidin II) #status experimental <KBDV>
 F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
 F:417-488/Region: glycine/histidine/lysine-rich
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
 F:27-591, 82-93, 106-125, 141-144, 205-217, 228-247, 263-266, 327-339, 350-369/Disulfide bond
 F:87, 168, 169, 204/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
 F:19/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:382/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:398, 406, 512/Binding site: carbohydrate (Ser) (covalent) #status experimental

F:399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status ex
F:498-499/Cleavage site: Arg-Thr (Kallikrein) #status experimental

Query Match 90.7%; Score 68; DB 1; Length 621;
Best Local Similarity 83.3%; Pred. No. 0.0026;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGHQOQHGLGHC 12
|||:|||||
Db 463 HGHQOQHGLGHC 474

RESULT 4

K-kininogen, HMW precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 30-Sep-1993
C:Accession: C25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family.
A:Reference number: A92625; MUID:87137443
A:Molecule type: DNA
A:Accession: C25486
A:Residues: 1-264 <KIT>
C:Comment: The nucleotide sequence was obtained from Genbank, release 55.0.
C:Superfamily: kininogen; cystatin homology

Query Match 82.7%; Score 62; DB 2; Length 264;
Best Local Similarity 75.0%; Pred. No. 0.0098;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 HGHQOQHGLGHC 12
|||:|||||
Db 75 HGHQOQHGLGHC 86

RESULT 5

kininogen, HMW I precursor - rat
N:Contains: Bradykinin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C:Accession: A25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family.
A:Reference number: A92625; MUID:87137443
A:Accession: A25486
A:Molecule type: mRNA
A:Residues: 1-639 <KIT>
A:Note: the authors translated the codon CAA for residue 347 as Asn
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-639/Product: kininogen, HMW I #status predicted <MAT>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>

Query Match 82.7%; Score 62; DB 2; Length 639;
Best Local Similarity 75.0%; Pred. No. 0.024;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 HGHQOQHGLGHC 12
|||:|||||
Db 450 HGHQOQHGLGHC 461

RESULT 6

C81428
peptidyl-prolyl cis-trans isomerase Cj0115 [imported] - Campylobacter jejuni (strain
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C:Accession: C81428
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chli
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912
A:Accession: C81428
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-189 <PAR>
A:Cross-references: GB:AL139074; GB:AL111168; NID:96967505; PIDN:CA872599.1; PID:9696
A:Experimental source: serotype O2, strain NCIC 11168
C:Genetics:
A:Gene: slyD; Cj0115

Query Match 73.3%; Score 55; DB 2; Length 189;
Best Local Similarity 66.7%; Pred. No. 0.09;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGHQOQHGLGHC 12
|||:|||||
Db 165 HGHQOQHGLGHC 176

RESULT 7

C27115
K-kininogen, LMW precursor - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
C:Accession: C27115; A25486
R:Fung, W.P.; Schreiber, G.
J. Biol. Chem. 262, 9298-9308, 1987
A:Title: Structure and expression of the genes for major acute phase alpha-1-protein
A:Reference number: A92653; MUID:87250580
A:Accession: C27115
A:Molecule type: DNA
A:Residues: 1-290 <FUN>
R:Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
J. Biol. Chem. 262, 2345-2351, 1987
A:Title: Differing utilization of homologous transcription initiation sites of rat K
A:Reference number: A25486; MUID:87137465
A:Accession: A25486
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-48 <KAG>
A:Cross-references: GB:J02662; NID:9205071; PIDN:AAA41483.1; PID:9205072
C:Superfamily: kininogen; cystatin homology
F:19-65/Domain: cystatin homology (fragment) <CY5>

Query Match 73.3%; Score 55; DB 2; Length 290;
Best Local Similarity 66.7%; Pred. No. 0.14;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 HGHQOQHGLGHC 12
|||:|||||
Db 124 HGHQOQHGLGHC 135

RESULT 8

A27115
major acute phase alpha-1 protein 1 - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A27115
R:Fung, W.P.; Schreiber, G.
J. Biol. Chem. 262, 9298-9308, 1987
A:Title: Structure and expression of the genes for major acute phase alpha-1-protein

A:Reference number: A92653; MUID:87250580
A:Accession: A27115
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-315 <FUN>
C:Genetics:
A:Gene: MAP1
C:Superfamily: Kininogen; cystatin homology
F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 73.3%; Score 55; DB 2; Length 315;
Best Local Similarity 66.7%; Pred. No. 0.15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 HGEQOHLGHC 12
DB 148 HGHQKPHGLGHC 159

RESULT 9

hypothetical protein SC5C7.34 SC5C7.34 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35241
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A:Reference number: 221572
A:Accession: T35241
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-314 <SEP>
A:Cross-references: EMBL:AL031515; PIDN:CAA20646.1; GSPDB:GN00070; SCOEDB:SC5C7.34
A:Experimental source: strain A5(2)
C:Genetics:
A:Gene: SCOEDB:SC5C7.34

Query Match 72.0%; Score 54; DB 2; Length 314;
Best Local Similarity 75.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGEQOHLGHC 12
DB 18 HGHQGHGHC 29

RESULT 10

hemolymph antifungal protein precursor - flesh fly (Sarcophaga peregrina)
C:Species: Sarcophaga peregrina
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Sep-1997
C:Accession: A45969
R:Iijima, R.; Kurata, S.; Natori, S.
J. Biol. Chem. 268, 12055-12061, 1993
A:Title: Purification, characterization, and cDNA cloning of an antifungal protein from
A:Reference number: A45969; MUID:93280179
A:Accession: A45969
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-85 <IID>
A:Cross-references: GB:D13797; NID:g391903; PID:d1003460; PID:g391904

Query Match 69.3%; Score 52; DB 2; Length 85;
Best Local Similarity 66.7%; Pred. No. 0.12;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGEQOHLGHC 12
DB 22 HGHQGHGHC 33

RESULT 11

probable zinc transporter At2g46800 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F19D11.8
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02681; D84907
R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.;
submitted to the EMBL Data Library, September 1998
A:Description: Arabidopsis thaliana chromosome II BAC F19D11 genomic sequence.
A:Reference number: T02681
A:Accession: T02681
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-398 <ROU>
A:Cross-references: EMBL:AC005310; NID:g3510247; PID:g3510254
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon,
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: D84907
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <STO>
A:Cross-references: GB:AB002093; NID:g3510254; PIDN:MAC3498.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g46800; F19D11.8
A:Map position: 2

Query Match 68.0%; Score 51; DB 2; Length 398;
Best Local Similarity 66.7%; Pred. No. 0.84;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGEQOHLGHC 12
DB 184 HGHSHGHC 195

RESULT 12

ATP/GTP-binding protein (ImpB/MucB/Samb family) BH2741 [imported] - Bacillus halodura
E83992
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C:Accession: E83992
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20263314
A:Accession: E83992
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <STO>
A:Cross-references: GB:AP001516; GB:BA000004; NID:910175192; PIDN:BAB06460.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2741

Query Match 68.0%; Score 51; DB 2; Length 409;
Best Local Similarity 58.3%; Pred. No. 0.87;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 HGEQOHLGHC 12
DB 242 HSHDKKKGHC 253

RESULT 13

T07618
 cold stress protein c11.8 - garden pea
 C:Species: Pisum sativum (garden pea)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
 C:Accession: T07618
 R:Kung, C.C.; Yeh, K.W.; Lin, C.Y.; Chen, Y.M.
 Bot. Bull. Acad. Sin. 39, 9-15, 1998
 A:Title: Characterization of a pea gene responsive to low temperature.
 A:Reference number: Z16055
 A:Accession: T07618
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-110 <RUN>
 A:Cross-references: EMBL:U24398; NID:g2947080; PID:g2947081
 A:Experimental source: cv. Taichung 9
 C:Comment: This protein is cold-induced.
 C:Superfamily: cold stress protein COR19

Query Match 67.3%; Score 50.5; DB 2; Length 110;
 Best Local Similarity 42.9%; Pred. No. 0.27;
 Matches 9; Conservative 2; Mismatches 1; Indels 9; Gaps 1;

OY 1 HGHE-----QHGGLHG 12
 |||||
 Db 33 HGHEHGGAETKGEHGFHG 53

RESULT 14

T48099
 hypothetical protein T20010.200 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48099
 R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeidler, K.; Mewes, H.W.; Rudd, S.; Lem
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24484
 A:Accession: T48099
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-199 <ORF>
 A:Cross-references: EMBL:AL163816
 A:Experimental source: cultivar Columbia; BAC clone T20010
 C:Genetics:
 A:Map position: 3
 A:Introns: 163/2
 A:Note: T20010.200

Query Match 66.7%; Score 50; DB 2; Length 199;
 Best Local Similarity 66.7%; Pred. No. 0.59;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGHEQOGLG 12
 |||||
 Db 92 HGHRGHRGHRG 103

RESULT 15

S66148
 gene pipsqueak protein A short form - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000
 C:Accession: S66148
 R:Weber, U.; Siegel, V.; Mlodzik, M.
 EMBO J. 14, 6247-6257, 1995
 A:Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for t
 A:Reference number: S66148; MUID:96134923
 A:Accession: S66148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-535 <WEB>
 A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CA62473.1; PID:g1149499

C:Genetics:
 A:Gene: pipsqueak
 C:Superfamily: POZ domain homology
 F:21-123/Domain: POZ domain homology <POZ>

Query Match 66.7%; Score 50; DB 2; Length 535;
 Best Local Similarity 66.7%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHEQOGLG 12
 |||||
 Db 332 HEHEHHNHGHRG 343

Search completed: July 6, 2001, 09:17:57
 Job time: 643 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:29 ; Search time 37.59 Seconds

(without alignments)
10.936 Million cell updates/sec

Title: US-09-437-912-1

Perfect score: 75
Sequence: 1 HGHGQGHG 12

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	75	100.0	644	1	KNG_HUMAN
2	68	90.7	619	1	KNH2_BOVIN
3	68	90.7	621	1	KNH1_BOVIN
4	62	82.7	639	1	KNG_RAT
5	62	82.7	661	1	KNG_MOUSE
6	56	74.7	693	1	CAUP_DROME
7	52	69.3	83	1	ANPE_SARPE
8	51	68.0	118	1	S109_RABIT
9	49	65.3	670	1	VG50_HSV1
10	49	65.3	2038	1	FSH_DROME
11	48	64.0	439	1	COT1_YEAST
12	47	62.7	515	1	KE4L_DROME
13	46	61.3	469	1	KE4_HUMAN
14	45	60.0	155	1	KE4_PIG
15	45	60.0	421	1	BR3A_MOUSE
16	45	60.0	503	1	ZNT1_MOUSE
17	45	60.0	507	1	ZNT1_RAT
18	43	57.3	335	1	HYPB_RHOCA
19	42	56.0	212	1	SLYD_AERAY
20	42	56.0	232	1	DHN3_PEA
21	42	56.0	258	1	BOX5_NOTVI
22	42	56.0	604	1	SR68_DROME
23	42	56.0	619	1	SR68_HUMAN
24	42	56.0	622	1	SR68_DROME
25	42	56.0	622	1	SR68_MOUSE
26	41.5	55.3	207	1	SR68_MOUSE
27	41.5	55.3	306	1	YQJ1_ECOLI
28	41.5	55.3	306	1	CH36_DROME
29	41	54.7	352	1	KE4_MOUSE
30	41	54.7	449	1	KE4_MOUSE
31	40.5	54.0	447	1	YHJ3_ARATH
32	40	53.3	335	1	PRSG_ECOLI
33	40	53.3	513	1	PIX1_DROME

ALIGNMENTS

RESULT 1	ID	KNG_HUMAN	STANDARD	PRT	644 AA
AC	P01042	P01043			
DT	21-JUL-1986	(Rel. 01, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].				
GN	KNG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).				
RC	TISSUE=Liver;				
RX	MEDLINE=85234582; PubMed=2989293;				
RA	Takagaki Y., Kitamura N., Nakanishi S.;				
RT	"Cloning and sequence analysis of cDNAs for human high molecular weight and low molecular weight prekallinogens. Primary structures of two human prekallinogens.";				
RL	J. Biol. Chem. 260:8601-8609(1985).				
RN	[2]				
RP	GENE STRUCTURE.				
RX	MEDLINE=85234583; PubMed=2989294;				
RA	Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T., Nakanishi S.;				
RT	"Structural organization of the human kininogen gene and a model for its evolution.";				
RL	J. Biol. Chem. 260:8610-8617(1985).				
RN	[3]				
RP	SEQUENCE OF 1-401 FROM N.A.				
RX	MEDLINE=85122621; PubMed=6441591;				
RA	Ohkubo I., Kurachi K., Takasawa T., Shokawa H., Sasaki M.;				
RT	"Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kininogen.";				
RL	Biochemistry 23:5691-5697(1984).				
RN	[4]				
RP	SEQUENCE OF 379-644.				
RX	MEDLINE=86030270; PubMed=4054110;				
RA	Lotspiech F., Kellermann J., Henschen A., Poertsch B., Mueller-Esterl W.;				
RT	"The amino acid sequence of the light chain of human high-molecular-mass kininogen.";				
RL	Eur. J. Biochem. 152:307-314(1985).				
RN	[5]				
RP	SEQUENCE OF 381-389.				
RX	MEDLINE=90255622; PubMed=4952632;				
RA	Pierce J.V.;				
RT	"Structural features of plasma kinins and kininogens.";				
RL	Fed. Proc. 27:52-57(1968).				
RN	[6]				
RP	DISULFIDE BONDS.				
RA	Sueyoshi T., Miyata T., Kato H., Iwanaga S.;				
RT	"Disulfide bonds in bovine HMW kininogens.";				

34	40	53.3	568	1	DISC_DROME
35	39.5	52.7	155	1	DH8_CRAPL
36	39	52.0	176	1	CON8_NEUCR
37	39	52.0	330	1	LAFU_VITPA
38	39	52.0	417	1	RNDP_HUMAN
39	39	52.0	419	1	RNDP_RAT
40	39	52.0	423	1	BR3A_HUMAN
41	39	52.0	462	1	HEMO_HUMAN
42	39	52.0	476	1	KE4_MOUSE
43	39	52.0	515	1	DRTS_CRIFA
44	39	52.0	527	1	HSF8_LYCPS
45	39	52.0	527	1	HSF8_LYCPS

Seikagaku 56:808-808(1984).

-1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2) HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO FACTOR XIII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5) LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD CLOTTING.

-1- SUBCELLULAR LOCATION: SECRETED.

-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE PRODUCED BY ALTERNATIVE SPLICING.

-1- TISSUE SPECIFICITY: PLASMA.

-1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

-1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: R02566; AAA35497.1; -

CC EMBL: M11437; AAB59550.1; -

CC EMBL: M11438; AAB59550.1; JOINED.

CC EMBL: M11521; AAB59550.1; JOINED.

CC EMBL: M11522; AAB59550.1; JOINED.

CC EMBL: M11523; AAB59550.1; JOINED.

CC EMBL: M11524; AAB59550.1; JOINED.

CC EMBL: M11525; AAB59550.1; JOINED.

CC EMBL: M11526; AAB59550.1; JOINED.

CC EMBL: M11527; AAB59550.1; JOINED.

CC EMBL: M11528; AAB59550.1; JOINED.

CC EMBL: M11437; AAB59551.1; -

CC EMBL: M11438; AAB59551.1; JOINED.

CC EMBL: M11521; AAB59551.1; JOINED.

CC EMBL: M11522; AAB59551.1; JOINED.

CC EMBL: M11523; AAB59551.1; JOINED.

CC EMBL: M11524; AAB59551.1; JOINED.

CC EMBL: M11525; AAB59551.1; JOINED.

CC EMBL: M11526; AAB59551.1; JOINED.

CC EMBL: M11527; AAB59551.1; JOINED.

CC EMBL: M11528; AAB59551.1; JOINED.

CC PIR: A01279; KGHUHL.

CC PIR: A25276; A25276.

CC PIR: A01280; KGHUHL.

CC PIR: B25276; B25276.

CC PIR: S02482; S02482.

CC SWISS-2DPAGE; P01043; HUMAN.

CC MIM: 228960; -

CC InterPro: IPR000010; -

CC InterPro: IPR002395; -

CC Pfam: PF00031; cystatin; 3.

CC PRINTS: PR00334; KININOGEN.

CC PROSITE: PS00287; CYSTATIN; 2.

CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator; Bradykinin; Blood coagulation; Inflammatory response; Signal.

CC Alternative splicing.

CC SIGNAL 1 18

CC CHAIN 19 644 KININOGEN.

CC CHAIN 19 380 KININOGEN HEAVY CHAIN.

CC PEPTIDE 381 389 BRADYKININ.

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD.RES	19	19	
FT	DISULFID	28	614	
FT	DISULFID	83	94	
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	
FT	CARBOHYD	169	169	
FT	CARBOHYD	205	205	
FT	CARBOHYD	294	294	
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPLIC	402	427	
FT	VARSPLIC	428	644	
FT	CONFLICT	593	593	
SO	SEQUENCE	644 AA;	71945 MW;	

Query Match 100.08; Score 75; DB 1; Length 644;
Best Local Similarity 100.08; Pred. No. 0.0001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 463 HGHEQHGHLGHC 474

1 HGHEQHGHLGHC 12
|||||||

RESULT 2
KMH2_BOVIN STANDARD; PRT; 619 AA.
AC P01045;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KININOGEN, HMW II PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84014106; PubMed=6571699;
RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT "A single gene for bovine high molecular weight and low molecular
RT weight kininogens.";
RL Nature 305:545-549(1983).
RP SEQUENCE OF 19-376.
RX MEDLINE=87137530; PubMed=3546295;
RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,

VSPPTSMAPADDEERDSKREGHTR -> SHLRSCCEKGR
PKKAAEPASEREVS (IN ISOFORM LMW).
MISSING (IN ISOFORM LMW).
T -> I (IN REF. 1).
3132B4CBAF8FB7E CRC64;

N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).

PYRROLIDONE CARBOXYLIC ACID.
INTERCHAIN.

RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion."; J. Biol. Chem. 262:2768-2779(1987).
 RN [3]
 RP SEQUENCE OF 376-391.
 RX MEDLINE-70180420; PubMed-4986212;
 RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 RT bonds and of methionyl bonds in kininogen-II."; J. Biochem. 67:313-323(1970).
 RN [4]
 RP SEQUENCE OF 387-455.
 RX MEDLINE-76260155; PubMed-956151;
 RA Han Y.N., Kato H., Iwanaga S., Suzuki T.;
 RT "Primary structure of bovine plasma high-molecular-weight kininogen.
 RT The amino acid sequence of a glycopeptide portion (fragment 1)
 RT following the C-terminus of the bradykinin moiety."; J. Biochem. 79:1201-1222(1976).
 RN [5]
 RP SEQUENCE OF 456-496.
 RX MEDLINE-75170265; PubMed-1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 RT kininogen. Amino acid sequence of a fragment ('histidine-rich
 RT peptide') released by plasma kallikrein."; J. Biochem. 77:55-68(1975).
 RN [6]
 RP FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C)
 CC NARURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 398.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTRAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC
 CC EMBL: V01492; CAA24736.1; -;
 CC EMBL: V01492; CAA24737.1; ALT_SEQ.
 CC PIR: A01283; KGRH02.
 CC PIR: B29559; B29559.
 CC HSSP: P04129; IAF1.
 CC InterPro: IPR000010; -;
 CC InterPro: IPR002395; -;
 CC Pfam: PF00031; cystatin.3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSTATIN; 2.
 CC Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 CC Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
 CC Inflammatory response.

FT SIGNAL 1 18
 FT CHAIN 19 619 KININOGEN, HMW II.
 FT CHAIN 19 376 HEAVY CHAIN.
 FT PEPTIDE 378 386 BRADYKININ.
 FT CHAIN 387 619 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 256 CYSTATIN-LIKE 2.
 FT DOMAIN 257 376 CYSTATIN-LIKE 3.
 FT MOD. RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 136 136 PARTIAL.
 FT CARBOHYD 168 168 OR 169.
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 400 400 INTERCHAIN.
 FT DISULFID 27 589
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 261 264
 FT DISULFID 325 337
 FT DISULFID 348 367
 FT VARIANT 398 398
 FT VARIANT 401 401
 FT VARIANT 454 454
 FT VARIANT 454 454
 SQ SEQUENCE 619 AA; 68710 MW; F04320A8BE0E0DA CRC64;
 T -> P.
 L -> V.
 H -> K.
 Query Match 90.7%; Score 68; DB 1; Length 619;
 Best Local Similarity 83.3%; Pred. No. 0.0012;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGHEOHGIGHG 12
 DB 461 HGHOHGHGHC 472
 RESULT 3
 KNL1_BOVIN
 ID KNL1_BOVIN STANDARD; PRT; 621 AA.
 AC P01044;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE KININOGEN, HMW I PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS:
 DE BRADYKININ].
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-84014106; PubMed-6571699;
 RA Kitamura N., Takagaki T., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
 RT "A single gene for bovine high molecular weight and low molecular
 RT weight kininogens."; J. Biol. Chem. 262:545-549(1983).
 RL Nature 305:545-549(1983).
 RN [2]
 RP SEQUENCE OF 19-378.
 RX MEDLINE-87137530; PubMed-3546295;
 RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
 RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion."; J. Biol. Chem. 262:2768-2779(1987).
 RN [3]
 RP SEQUENCE OF 378-393.
 RX MEDLINE-70180420; PubMed-4986212;

RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [4]
 RP SEQUENCE OF 458-498.
 RX MEDLINE=75170265; PubMed=1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 kininogen. Amino acid sequence of a fragment ('histidine-rich
 peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
 CC NATRIURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4E) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 400.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: V01491; CAA24735.1; -
 CC PIR: A01281; KGB0H1.
 CC PIR: A29559; A29559.
 CC InterPro: IPR000010; -
 CC InterPro: IPR002395; -
 CC DR Pfam: PF00031; cystatin; 3.
 CC DR PRINTS: PR00334; KININOGEN.
 CC DR PROSITE: PS00287; CYSSTATIN; 2.
 CC KW Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 KW Thiol protease inhibitor; Bradykinin; Blood coagulation;
 KW Inflammatory response; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 621
 FT CHAIN 19 378
 FT PEPTIDE 380 388
 FT CHAIN 389 621
 FT DOMAIN 19 135
 FT DOMAIN 136 257
 FT DOMAIN 258 378
 FT MOD_RES 19 19
 FT CARBOHYD 87 87
 FT CARBOHYD 136 136
 FT CARBOHYD 168 168
 FT CARBOHYD 167 197
 FT CARBOHYD 204 204
 FT DISULFID 27 591
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144

FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;
 Query Match 90.7%; Score 68; DB 1; Length 621;
 Best Local Similarity 83.3%; Pred. No. 0.0013;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGHEQOHGLGHC 12
 DB 463 HGHEQOHGLGHC 474
 RESULT 4
 ID KNG_RAT
 AC P08934; P08933; STANDARD; PRT; 639 AA.
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
 OS KNG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RX MEDLINE=87137443; PubMed=3029068;
 RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
 RT "Differing expression patterns and evolution of the rat kininogen
 RT gene family.";
 RL J. Biol. Chem. 262:2190-2198(1987).
 RN [2]
 RP SEQUENCE FROM N.A. (LMW ISOFORM).
 RX MEDLINE=86008264; PubMed=2413018;
 RA Furuto-kato S., Matsumoto A., Kitamura N., Nakanishi S.;
 RT "Primary structures of the mRNAs encoding the rat precursors for
 RT bradykinin and T-kinin. Structural relationship of kininogens with
 RT major acute phase protein and alpha 1-cysteine proteinase
 RT inhibitor.";
 RL J. Biol. Chem. 260:12054-12059(1985).
 RN [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=BUFPALO;
 RX MEDLINE=87250580; PubMed=2439509;
 RA Fung W.-P., Schreiber G.;
 RT "Structure and expression of the genes for major acute phase alpha 1-
 RT protein (thioesterin) and kininogen in the rat.";
 RL J. Biol. Chem. 262:9298-9308(1987).
 RN [4]
 RP SEQUENCE OF 1-41 FROM N.A.
 RC STRAIN=WISTAR; TISSUE=Liver;
 RX MEDLINE=87137465; PubMed=3818598;
 RA Kagayama R., Kitamura N., Ohkubo H., Nakanishi S.;
 RT "Differing utilization of homologous transactivation initiation sites
 RT of rat K and T kininogen genes under inflammation condition.";
 RL J. Biol. Chem. 262:2345-2351(1987).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF

CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIN-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- MISCELLANEOUS: RAY EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
 CC HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
 CC T-II.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U29428; AAA41486.1; -;
 CC EMBL: M1884; AAA41487.1; -;
 CC EMBL: M14369; AAA41484.1; -;
 CC EMBL: M14365; AAA41485.1; ALT-SEQ.
 CC EMBL: M16455; AAA41482.1; -;
 CC PIR: A25486; A25486.
 CC PIR: A28055; A28055.
 CC InterPro: IPR000010; -;
 CC InterPro: IPR002395; -;
 CC Pfam: PF000334; cystatin. 3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSTATIN. 2.
 CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 CC Bradykinin; Blood coagulation; Inflammatory response; Signal;
 CC Alternative splicing; Multigene family.
 CC SIGNAL 1 18
 CC FT CHAIN 19 639 KININOGEN, HEAVY CHAIN.
 CC FT CHAIN 19 380 BRADYKININ.
 CC FT PEPTIDE 381 389 KININOGEN LIGHT CHAIN.
 CC FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
 CC FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 CC FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 CC FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 CC FT DOMAIN 439 514 HIS-RICH.
 CC FT DOMAIN 28 609 INTERCHAIN (BY SIMILARITY).
 CC FT DISULFID 83 94 BY SIMILARITY.
 CC FT DISULFID 107 126 BY SIMILARITY.
 CC FT DISULFID 142 145 BY SIMILARITY.
 CC FT DISULFID 206 218 BY SIMILARITY.
 CC FT DISULFID 229 248 BY SIMILARITY.
 CC FT DISULFID 264 267 BY SIMILARITY.
 CC FT DISULFID 328 340 BY SIMILARITY.
 CC FT DISULFID 351 370 BY SIMILARITY.
 CC FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 402 433 VSPSTIARVQERDRCNGQPIHGGMHAKO > RLINS
 CC FT VARSPLIC CEFKRLKAGAGPAPEKQAEASTVTP (IN ISOFORM
 CC FT VARSPLIC LMW).
 CC FT VARSPLIC 434 639 MISSING (IN ISOFORM LMW).
 CC FT CONFLICT 61 E-> K (IN REF. 2).
 CC FT SEQUENCE 639 AA: 70933 MW: D3172DF94FF56AF5 CRC64;
 CC
 CC Query Match 82.7%; Score 62; DB 1; Length 639;
 CC Best Local Similarity 75.0%; Pred. No. 0.01;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 HGHEOHGCHG 12
 |||||
 Db 450 KGHKPHGCHG 461
 RESULT 5
 KNG_MOUSE
 ID KNG_MOUSE STANDARD; PRT; 661 AA.
 AC 008677; 008676;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
 GN KNG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RC STRAIN=C57BL/6 x CBA; TISSUE=liver;
 RA Takano M., Kondoh J., Yajima K., Okamoto H.;
 RT "Molecular cloning of cDNAs for mouse low- and high- molecular
 RT kinogen".
 RT Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIN-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D84435; BAA19743.1; -;
 CC EMBL: D84415; BAA19742.1; -;
 CC MGD: MGI:1097705; Kng.
 CC DR MGD: MGI:1097705; Kng.
 CC DR InterPro: IPR000010; -;
 CC DR InterPro: IPR002395; -;
 CC DR InterPro: IPR003243; -;
 CC DR Pfam: PF000334; cystatin. 3.
 CC DR PRINTS: PR00334; KININOGEN.
 CC DR PROSITE: PS00287; CYSTATIN. 1.
 CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 CC Bradykinin; Blood coagulation; Inflammatory response; Signal;
 CC Alternative splicing.
 CC SIGNAL 1 18 POTENTIAL.
 CC FT CHAIN 19 661 KININOGEN.

```

FT CHAIN 19 379 KININOGEN HEAVY CHAIN.
FT CHAIN 380 388 BRADYKININ.
FT CHAIN 389 661 KININOGEN LIGHT CHAIN.
FT DOMAIN 136 135 CYSTATIN-LIKE 1.
FT DOMAIN 136 257 CYSTATIN-LIKE 2.
FT DOMAIN 258 379 CYSTATIN-LIKE 3.
FT DOMAIN 439 524 HIS-RICH.
FT DISULFID 28 631 INTERCHAIN (BY SIMILARITY).
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 107 125 BY SIMILARITY.
FT DISULFID 141 144 BY SIMILARITY.
FT DISULFID 205 217 BY SIMILARITY.
FT DISULFID 228 247 BY SIMILARITY.
FT DISULFID 263 266 BY SIMILARITY.
FT DISULFID 327 339 BY SIMILARITY.
FT DISULFID 350 369 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 401 432 VSPPLYAROEERDAETEGPHGHGMLHEKO -> RLTRA
FT VARSPLIC 401 432 CEKGRLSKAGAEPAERQAESOVYQ (IN ISOFORM
FT VARSPLIC 401 432 LMW).
FT VARSPLIC 401 432 MISSING (IN ISOFORM LMW).
SQ SEQUENCE 661 AA: 73102 MW: 774460258058796E CRC64;

```

Query Match
Best Local Similarity 82.7%; Score 62; DB 1; Length 661;
Matches 9; Conservative 75.0%; Pred. No. 0.012; Mismatches 1; Indels 0; Gaps 0;

OY 1 HGHEQHGGLG 12
Db 460 HGHGKHGGLG 471

```

RESULT 6
CAUP_DROME STANDARD; PRT; 693 AA.
AC P54269;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HOMEOBOX PROTEIN CAUPOLICAN.
GN CAUP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96180722; PubMed=8620542;
RA Gomez-Skarmeta J.-L., del Corral R.D., de la Calle-Mustienes E.,
RA Ferrer-Marco D., Modolell J.;
RA Araucan and caupolican, two members of the novel iroquois complex,
RT encode homeoproteins that control proneural and vein-forming genes.;
RT Cell 85:95-110(1996).
RL
CC -1- FUNCTION: CONTROLS PRONEURAL AND VEIN FORMING GENES. POSITIVE
CC TRANSCRIPTIONAL CONTROLLER OF AC-SC (ACHAETE-SCUTE). MAY ACT AS AN
CC ACTIVATOR THAT INTERACTS WITH THE TRANSCRIPTIONAL COMPLEX
CC ASSEMBLED ON THE AC AND SC PROMOTERS AND PARTICIPATES IN
CC TRANSCRIPTION INITIATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE TALE/INO FAMILY OF HOMEOBOX PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: X95178; CAA64485.1; -.
DR HSSP: P02833; ISAN.
DR FlyBase: FBgn0015919; caup.
DR InterPro: IPR001356; -.
DR Pfam: PF00046; homeobox; 1.
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS00071; HOMEOBOX_2; 1.
KW Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
KW Developmental protein.
FT DNA_BIND 226 288 HOMEOBOX (TALE-TYPE).
FT DOMAIN 300 303 POLY-ASP.
FT DOMAIN 405 418 POLY-GLN.
FT DOMAIN 501 516 POLY-GLN.
FT DOMAIN 517 528 POLY-HIS.
FT DOMAIN 565 572 POLY-SER.
FT DOMAIN 613 624 POLY-SER.
SQ SEQUENCE 693 AA: 73749 MW: 860064369C6D619 CRC64;

```

Query Match
Best Local Similarity 74.7%; Score 56; DB 1; Length 693;
Matches 9; Conservative 75.0%; Pred. No. 0.11; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGHEQHGGLG 12
Db 656 HGHGKHGGLG 667

```

RESULT 7
ANTF_SARPE STANDARD; PRT; 85 AA.
AC 008617;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ANTI-FUNGAL PROTEIN PRECURSOR (AFP).
OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestrioidae; Sarcophagidae; Sarcophaga.
OX NCBI_Taxid=7386;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-54.
RX TISSUE=Fat body;
RX MEDLINE=93280179; PubMed=8505329;
RA Iijima R., Kurata S., Natori S.;
RT Purification, characterization, and cDNA cloning of an antifungal
RT protein from the hemolymph of Sarcophaga peregrina (Flesh fly)
RT larvae.;
RT J. Biol. Chem. 268:12055-12061(1993).
CC -1- FUNCTION: THIS PROTEIN INHIBITS THE GROWTH OF A VARIETY OF
CC FUNGAL SPECIES. THE ANTI-FUNGAL ACTIVITY OF THIS PROTEIN IS
CC ENHANCED BY THE PRESENCE OF SARCOTOXIN IA.
CC -1- SUBUNIT: HOMODIMER.
CC -1- TISSUE SPECIFICITY: HEMOLYMPH.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D13797; BAA02954.1; -.
DR PIR: A45969; A45969.
CC Functid: Signal; Repeat.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 85 ANTI-FUNGAL PROTEIN.
FT DOMAIN 19 73 2 X 7 AA REPEATS OF Q-H-G-H-G-Q.
FT REPEAT 19 25 1.

```


Query Match 65.3%; Score 49; DB 1; Length 670;
 Best Local Similarity 66.7%; Pred. No. 1.3;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHGQHGGLGHC 12
 ||| ||| |||
 DB 637 HGHGHHGHGHC 648

RESULT 10

FSH_DROME
 ID FSH_DROME STANDARD; PRT; 2038 AA.
 AC P13709; P13710;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE FEMALE STERILE HOMEOTIC PROTEIN (FRAGILE-CHORION MEMBRANE PROTEIN).
 GN FSH OR FSH.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89276730; PubMed=2567251;
 RA Hayes S.R., Mozer B.A., Bhatia-Dey N., David I.B.;
 RT "The Drosophila fish locus, a maternal effect homeotic gene, encodes
 RT apparent membrane proteins";
 RL Dev. Biol. 134:246-257(1989).
 CC -!- FUNCTION: REQUIRED MATERIALLY FOR PROPER EXPRESSION OF OTHER
 CC HOMEOTIC GENES INVOLVED IN PATTERN FORMATION, SUCH AS UBX.
 CC -!- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.
 CC -!- SIMILARITY: CONTAINS 2 BROMODOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 ET DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M23221; AAA28540.1; -
 CC EMBL: M23222; AAA28541.1; ALT_TERM.
 CC EMBL: M15762; AAA70424.1; -
 CC EMBL: M15763; AAA70423.1; -
 CC EMBL: M15764; AAA70422.1; -
 CC PIR: A43742; A43742.
 CC DR HSSP: P04002; IMFA.
 CC DR FLYBASE: FBgn0004656; fs(1)h.
 CC DR InterPro: IPR001487; -
 CC DR Pfam: PF00439; bromodomain; 2.
 CC DR PRINTS: PR00503; BROMODOMAIN.
 CC DR PROSITE: PS00633; BROMODOMAIN_1; 2.
 CC DR PROSITE: PS50014; BROMODOMAIN_2; 2.
 CC KW Developmental protein; Bromodomain; Transmembrane; Repeat.
 FT DOMAIN 51 123 BROMODOMAIN 1.
 FT DOMAIN 495 567 BROMODOMAIN 2.
 FT DOMAIN 945 1106 ET DOMAIN.
 FT TRANSMEM 330 350 POTENTIAL.
 FT TRANSMEM 451 471 POTENTIAL.
 FT TRANSMEM 750 770 POTENTIAL.
 FT TRANSMEM 790 810 POTENTIAL.
 FT TRANSMEM 816 830 POTENTIAL.
 FT TRANSMEM 874 894 POTENTIAL.
 FT TRANSMEM 1731 1751 POTENTIAL.
 FT TRANSMEM 1939 1959 POTENTIAL.
 FT VARIANT 909 909 G -> A.
 FT VARIANT 1022 1022 H -> RRPYY.
 FT SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match 65.3%; Score 49; DB 1; Length 2038;
 Best Local Similarity 66.7%; Pred. No. 4.1;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHGQHGGLGHC 12
 ||| ||| |||
 DB 596 HGHGHHGHGHC 607

RESULT 11

COT1_YEAST
 ID COT1_YEAST STANDARD; PRT; 439 AA.
 AC P32798;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COBALT UPTAKE PROTEIN COT1.
 GN COT1 OR YOR316C OR O6131.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92375034; PubMed=1508175;
 RA Conklin D.S., McMaster J.A., Culbertson M.R., Kung C.;
 RT "COT1, a gene involved in cobalt accumulation in Saccharomyces
 RT cerevisiae.";
 RL Mol. Cell. Biol. 12:3678-3688(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=97051589; PubMed=8896266;
 RA Pearson B.M., Hernando Y., Payne J., Wolf S.S., Kalogeropoulos A.,
 RA Schweizer M.;
 RT "Sequencing of a 35.71 kb DNA segment on the right arm of yeast
 RT chromosome XV reveals regions of similarity to chromosomes I and
 RT XIII.";
 RL Yeast 12:1021-1031(1996).
 CC -!- FUNCTION: PROBABLY RESPONSIBLE FOR THE UPTAKE OF COBALT IONS. IT
 CC APPEARS TO ACT IN A DOSAGE-DEPENDENT MANNER TO COUNTERACT THE
 CC ADVERSE EFFECTS OF COBALT IONS ON CELLS. IT MAY PARTICIPATE IN
 CC THE REGULATION OF COBALT LEVELS UNDER NORMAL PHYSIOLOGICAL
 CC CONDITIONS AND MAY BE IMPORTANT IN THE SUPPLY OF METAL THAT IS
 CC REQUIRED FOR METALLOENZYME OR COFACTOR SYNTHESIS. IT REDUCES THE
 CC TOXICITY OF COBALT AND RHODIUM IONS. OTHER COMPONENTS RESPONSIBLE
 CC FOR COBALT TRANSPORT EXIST.
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MEMBRANE. ANOTHER POSSIBILITY
 CC EXISTS THAT IT IS ASSOCIATED WITH ANOTHER UNIDENTIFIED MEMBRANE
 CC THAT HAS BEEN ENRICHED IN THE MITOCHONDRIAL MEMBRANE FRACTIONS.
 CC -!- SIMILARITY: BELONGS TO THE SLC30A FAMILY OF TRANSPORTERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M88252; AAA74884.1; -
 CC EMBL: X90565; CAA62171.1; -
 CC EMBL: Z75224; CAA99636.1; -
 CC PIR: S31302; S31302.
 CC SGD: S0005843; COT1.
 CC DR InterPro: IPR002524; -
 CC DR Pfam: PF01545; Cation_efflux; 1.
 CC DR Transport; Cobalt; Mitochondrion; Transmembrane.
 FT TRANSMEM 10 27 POTENTIAL.
 FT TRANSMEM 43 60 POTENTIAL.
 FT TRANSMEM 78 100 POTENTIAL.

Query Match 60.0%; Score 45; DB 1; Length 421;
Best Local Similarity 63.6%; Pred No. 3.6;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 1 HGHEDQHGIGH 11
| | | | |
| | | | |
Db 190 HPHPHMHGIGH 200

Search completed: July 6, 2001, 09:26:36
Job time: 967 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:24 ; Search time 118.42 Seconds
(without alignments)
13.407 Million cell updates/sec

Title: US-09-437-912-1
Perfect score: 75
Sequence: 1 HGHQDQHGIGHG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	82.7	126	11 009016	009016 rattus norv
2	56	74.7	693	5 09VU00	09VU00 drosophila
3	55	73.3	189	2 09P112	09P112 campylobact
4	54	72.0	119	5 08VP13	08VP13 drosophila
5	54	72.0	314	2 086731	086731 streptomyc
6	54	72.0	511	5 09VXG3	09VXG3 drosophila
7	51	68.0	79	10 09M435	09M435 quercus rob
8	51	68.0	86	10 022671	022671 alnus glutl
9	51	68.0	99	10 09ZRC7	09ZRC7 alnus glutl
10	51	68.0	348	5 09W2X1	09W2X1 drosophila
11	51	68.0	398	10 081036	081036 arabisidops
12	51	68.0	409	2 09K9A8	09K9A8 bacillus ha
13	51	68.0	457	5 09W4I6	09W4I6 drosophila
14	51	68.0	605	5 077280	077280 drosophila
15	51	68.0	989	5 09W2S4	09W2S4 drosophila
16	50.5	67.3	110	10 064396	064396 pisum sativ
17	50	66.7	199	10 09LYB2	09LYB2 arabisidops
18	50	66.7	213	5 09GTN0	09GTN0 drosophila
19	50	66.7	1064	5 09V5N1	09V5N1 drosophila

20	50	66.7	1085	5 024455	024455 drosophila
21	49.5	66.0	218	5 09V3P9	09V3P9 drosophila
22	49	65.3	206	5 09GTN1	09GTN1 drosophila
23	49	65.3	385	5 09VWX5	09VWX5 drosophila
24	49	65.3	450	5 027920	027920 bradyzia hy
25	49	65.3	457	5 026227	026227 rhyndoscia
26	49	65.3	554	5 09W4C1	09W4C1 drosophila
27	49	65.3	686	5 09VWS0	09VWS0 drosophila
28	49	65.3	735	5 09NES7	09NES7 caenorhabdi
29	49	65.3	1937	5 09W3I3	09W3I3 drosophila
30	48	64.0	2262	5 09Y0E4	09Y0E4 drosophila
31	48	64.0	2262	5 09W0Q4	09W0Q4 drosophila
32	47	62.7	160	10 043520	043520 lycopersico
33	47	62.7	179	10 09M0L8	09M0L8 arabisidops
34	47	62.7	277	10 049678	049678 arabisidops
35	47	62.7	490	2 09X9W6	09X9W6 streptomyc
36	46.5	62.0	390	5 018401	018401 caenorhabdi
37	46	61.3	168	5 09VWM5	09VWM5 drosophila
38	46	61.3	198	5 09NNV9	09NNV9 plasmodium
39	46	61.3	505	5 09V3C6	09V3C6 drosophila
40	46	61.3	519	5 09VU19	09VU19 drosophila
41	45	60.0	317	2 005673	005673 mycobacteri
42	45	60.0	420	11 09QVZ9	09QVZ9 mus sp. brn
43	45	60.0	492	11 09JKN2	09JKN2 mus musculi
44	45	60.0	494	5 09GRW9	09GRW9 drosophila
45	45	60.0	495	5 09VEX1	09VEX1 drosophila

ALIGNMENTS

RESULT	2	PRELIMINARY:	PRT:	693 AA.
ID	Q9VU00			
QY	1 HGHQDQHGIGHG 12			
DB	38 HGHQKPHGIGHG 49			
DT	01-JUL-1997 (TREMBLrel. 04, Created)			
DT	01-OCT-2000 (TREMBLrel. 04, Last sequence update)			
DE	K-KININOGEN (FRAGMENT).			
GN	KINR.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=DOMRVO;			
RX	MEDLINE=97468288; PubMed=9321484;			
RA	Harris E.L., Grigor M.R., Innes B.A., Harrop S.B., Kolke G.,			
RA	Jacob H.J.;			
RT	*Strain-specific deletions in exon 10 of rat K-kininogen and 11-			
RT	kininogen genes allow mapping of both genes to rat chromosome 11.*;			
RL	Mamm. Genome 8:791-792(1997).			
DR	EMBL: AF003623; AAC09070.1; -.			
DR	InterPro: IPR002395; -.			
DR	PRINTS: PR00334; KININOGEN.			
FT	NON_TER			
FT	NON_TER			
SO	SEQUENCE			
Query Match	Best Local Similarity	82.7%;	Score 62;	DB 11;
Matches	9;	Conservative	2;	Mismatches
			1;	Indels
				0;
				Gaps
				0;

AC Q9VU00; 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CAUP PROTEIN.
 GN CAUP OR CG10605.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayri A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegyan C.,
 RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson R., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER HOMEOBOX DOMAINS.
 DR EMBL: AE003540; AAF49895.1; -
 DR Flybase: FBgn0015919; caup.
 DR InterPro: IPR001356; -
 DR Pfam: PF00046; homeobox.1.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS00071; HOMEOBOX_2; 1.
 DR SMART: SM00389; HOX; 1.
 DR SMART: SM00389; HOX; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SO SEQUENCE 693 AA; 73667 MW; FBBR1616493F7EC9 CRC64;

RESULT 3
 ID Q9PJ12 PRELIMINARY; PRT; 189 AA.
 AC Q9PJ12;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE.
 GN SLYD OR CJO115.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parthill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagsels K., Karsley A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 DR EMBL: AL139074; CAB72599.1; -
 DR InterPro: IPR001179; -
 DR PROSITE: PS00059; FKBP_PP1ASE_3; 2.
 SO SEQUENCE 189 AA; 20132 MW; 47B5F5D047549D7F CRC64;

Query Match 73.3%; Score 55; DB 2; Length 189;
 Best Local Similarity 66.7%; Pred. No. 0.13;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 HGHEOQHGLGHC 12
 DB 165 HGHHHHHGHC 176
 RESULT 4
 ID Q9VYP3 PRELIMINARY; PRT; 119 AA.
 AC Q9VYP3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CG15733 PROTEIN.
 GN CG15733.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayri A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

CC ONE BINDING SITE FOR CALCIUM AND PHOSPHOLIPID (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ANNEXIN FAMILY.
 DR EMBL: AEO03502; AAF48609.1; -
 DR HSSP: P79134; IAVC.
 DR FlyBase: FBgn0030749; Anxb11.
 DR InterPro: IPR000130; -
 DR InterPro: IPR001464; -
 DR InterPro: IPR002395; -
 DR InterPro: IPR002965; -
 DR Pfam: PF00191; annexin.4.
 DR PRINTS: PRO0196; ANNEXIN.
 DR PRINTS: PRO0334; KININOGEN.
 DR PRINTS: PRO1217; PRICHEXTENSN.
 DR PROSITE: PS00223; ANNEXIN.3.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN.1.
 DR Annexin: Calcium/phospholipid-binding; Repeat.
 KW SEQUENCE 511 AA; 56214 MW; 7471286FC54283B CRC64;

Query Match 72.0%; Score 54; DB 5; Length 511;
 Best Local Similarity 75.0%; Pred. No. 0.53;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGHEQOHGICGHG 12
 ||| ||| |||
 DB 147 HGHGQGHGHC 158

RESULT 7
 ID 09M435 PRELIMINARY; PRT; 79 AA.
 AC 09M435;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PHASE-CHANGE RELATED PROTEIN PRECURSOR.
 OS Quercus robur (English oak).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fagales; Fagaceae; Quercus.
 OX NCBI_TaxID=38942;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-IN VITRO SHOOT CULTURES;
 RA Gil B., Pastoriza E.M., Ballester A., Sanchez C.;
 RT "Identification of a phase-change related mRNA in oak shoot cultures
 derived from basal sprouts and crown branches.",
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ271778; CAB72442.1; -
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 79 AA; 8414 MW; 8E45CABF40F00B6F CRC64;

Query Match 68.0%; Score 51; DB 10; Length 79;
 Best Local Similarity 66.7%; Pred. No. 0.24;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHEQOHGICGHG 12
 ||| ||| |||
 DB 44 HGHGQGHGHC 55

RESULT 8
 ID 022671 PRELIMINARY; PRT; 86 AA.
 AC 022671;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE AG164 PROTEIN PRECURSOR.
 GN AG164.
 OS Alnus glutinosa (Alder).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97348585; PubMed-9204569;
 RA Pawlowski K., Twigg P., Dobritsa S., Guan C., Mullin B.C.;
 RT "A nodule-specific gene family from Alnus glutinosa encodes glycine-
 and histidine-rich proteins expressed in the early stages of
 actinorhizal nodule development.";
 RL Mol. Plant Microbe Interact. 10:656-664(1997).
 DR EMBL: Y08436; CAA69708.1; -
 KW SIGNAL.
 FT SIGNAL.
 FT CHAIN 30 86 POTENTIAL.
 SQ SEQUENCE 86 AA; 9188 MW; DB5B7EF88C8899A CRC64;

Query Match 68.0%; Score 51; DB 10; Length 86;
 Best Local Similarity 66.7%; Pred. No. 0.26;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHEQOHGICGHG 12
 ||| ||| |||
 DB 50 HGHVHGHGHC 61

RESULT 9
 ID 09ZRC7 PRELIMINARY; PRT; 99 AA.
 AC 09ZRC7;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE ACTINORHIZAL NODULIN AGNOD-GHRP.
 GN AGN784.
 OS Alnus glutinosa (Alder).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Dobritsa S.V., Mullin B.C.;
 RT "In vitro expression of actinorhizal nodulin AgNOD-GHRP and
 demonstration of its toxicity to Escherichia coli.";
 RL (in) Stacey G., Mullin B.C., Gresshoff P.M. (eds.);
 RL The Biology of Plant-Microbe Interactions:
 RL Proceedings of the 8th International Symposium on Molecular
 RL Plant-Microbe Interactions, pp.1-1, Unknown Publisher (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Twigg P.G.;
 RT "Isolation of a nodule-specific cDNA encoding a putative glycine-rich
 protein from Alnus glutinosa.";
 RL Thesis (1993), The University of Tennessee, Knoxville, TN, USA.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Pawlowski K., Twigg P.G., Dobritsa S.V., Guan C., Mullin B.C.;
 RT "A nodule-specific gene family from Alnus glutinosa encodes glycine
 and histidine-rich proteins expressed in the early stages of
 actinorhizal nodule development.";
 RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U69156; AAD00171.1; -
 DR InterPro: IPR002395; -
 DR PRINTS: PRO0334; KININOGEN.
 SQ SEQUENCE 99 AA; 10567 MW; 2ACBEAD57C070E83 CRC64;

Query Match Best Local Similarity 68.0%; Score 51; DB 10; Length 99;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHEQOHLGHC 12
Db 50 HGHRVHGHC 61

RESULT 10

O9W2X1 PRELIMINARY; PRT; 348 AA.
AC O9W2X1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG3961 PROTEIN.
GN CG3961.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amannatidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Fiankoch C., Baldwin D.,
RA Balder R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iobagwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AEO03449; AAF46566.1;
DR FlyBase; FBgn0030187; CG2961.
DR InterPro; IPR002395;
DR PRINTS; PR00334; KININGEN.
SO SEQUENCE 348 AA; 33212 MW; 8DB719D2AEBB374 CRC64;

Query Match 68.0%; Score 51; DB 5; Length 348;

Best Local Similarity 66.7%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGHEQOHLGHC 12
Db 74 HGTAQHGHC 85

RESULT 11

O81036 PRELIMINARY; PRT; 398 AA.
AC O81036;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE PUTATIVE ZINC TRANSPORTER.
GN F19D11.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Romling C.M., Fraser C.M., Somerville C.R., Venter J.C.,
RT "Arabidopsis thaliana chromosome II BAC F19D11 genomic sequence."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005310; AAC33498.1;
DR InterPro; IPR002524;
DR Pfam; PF01545; Cation_efflux; 1.
SO SEQUENCE 398 AA; 43827 MW; 7E20E0B2937BB23 CRC64;

Query Match Best Local Similarity 68.0%; Score 51; DB 10; Length 398;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHEQOHLGHC 12
Db 184 HGSHGHC 195

RESULT 12

O9K9A8 PRELIMINARY; PRT; 409 AA.
AC O9K9A8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ATP/GTP-BINDING PROTEIN (IMPB/MCB/SMB FAMILY).
GN BH2741.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=8665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RA Takami H., Nakasone K., Takaki Y.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001516; BAB06460.1;
DR InterPro; IPR001126;
DR Pfam; PF00817; IMS; 1.
SO SEQUENCE 409 AA; 46251 MW; 5721AC1D8FDD3722 CRC64;

Query Match 68.0%; Score 51; DB 2; Length 409;
Best Local Similarity 58.3%; Pred. No. 1.3;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 HGHEQOHLGHC 12

DB 242 HSHDKKQIGHG 253

```

RESULT 13
ID 09W416 PRELIMINARY; PRT: 457 AA.
AC 09W416;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE EG:84H4.4 PROTEIN.
GN EG:84H4.4 OR CG3081.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chame M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Ababayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borotova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegyan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasmo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., Mpherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relier K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003431; AAF45965.1;
DR FLYbase: FBgn0025613; EG:84H4.4.
SQ SEQUENCE 457 AA; 48919 MW; 70B4B7ADDD02E0AD CRC64;

```

Query Match 68.0%; Score 51; DB 5; Length 457;
 Best Local Similarity 66.7%; Pred. No. 1.4;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHEDOHGIGHG 12
 ||| |||
 DB 341 HGHNHGHGIGHG 352

```

RESULT 14
ID 077280 PRELIMINARY; PRT: 605 AA.
AC 077280;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE EG:84H4.4 PROTEIN.
GN EG:84H4.4 OR CG3081.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Ferraz C., Vidal S., Brun C., Bucheton A., Demaille J.G.;
RT "Sequencing the distal X chromosome of Drosophila melanogaster."
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL031766; CAA21135.1;
DR FLYbase: FBgn0025613; EG:84H4.4.
SQ SEQUENCE 605 AA; 64947 MW; B06C84ACAD7D2C84 CRC64;

```

Query Match 68.0%; Score 51; DB 5; Length 605;
 Best Local Similarity 66.7%; Pred. No. 1.9;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHEDOHGIGHG 12
 ||| |||
 DB 341 HGHNHGHGIGHG 352

```

RESULT 15
ID 09W254 PRELIMINARY; PRT: 989 AA.
AC 09W254;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG9732. PROTEIN.
GN CG9732.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chame M., Pfeiffer B.D.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Ababayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borotova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

SO SEQUENCE 469 AA: 50118 MW: 6504A1EF5AA6A5B9 CRC64:

Query Match 61.3%; Score 46; DB 1; Length 469;

Best Local Similarity 63.6%; Pred. No. 2.8;

Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHGEOHGLGH 11
111 111
Db 56 HGHSHHGHGH 66

RESULT 14

KE4_PIG STANDARD: PRT: 155 AA.

AC 029175: O9XT01; Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE HISTIDINE-RICH PROTEIN KE4 (FRAGMENTS).

GN HKEA.
OS Sus. scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE OF 1-124 FROM N.A.

RC TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Winfree A.K., Fredholm M., Davies W.;
RT "Evolution and characterization of a porcine small intestine cDNA
RT library: analysis of 839 clones.";
RL Mann. Genome 7:509-517(1996).

RP SEQUENCE OF 125-155 FROM N.A.
RC STRAIN=Belgian Landrace;
RA Chardon P., Rogel-Gaillard C., Peelman L.J., Verle M., Renard C.,
RA Vainan M.;

RT "Physical organization of the swine major histocompatibility complex
RT class II region.";
RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: F14787: CAA23256.1; -
DR EMBL: AF146397; AAD44801.1; -
DR EMBL: Transmembrane.

FT NON_TER 1
FT TRANSMEM 19 39 POTENTIAL.
FT DOMAIN 49 59 HIS-RICH.
FT DOMAIN 81 86 POLY-GLU.
FT NON_CONS 124 125
FT NON_TER 155 155

SO SEQUENCE 155 AA: 16840 MW: 56354DE3CAAE4524 CRC64;

Query Match 60.0%; Score 45; DB 1; Length 155;

Best Local Similarity 58.3%; Pred. No. 1.3;

Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HGHGEOHGLGH 12
111 111
Db 49 HGHSHHGHGH 60

RESULT 15

BR3A_MOUSE STANDARD: PRT: 421 AA.

AC P17208; Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last sequence update)

DE BRAIN-SPECIFIC HOMEOBOX/POU DOMAIN PROTEIN 3A (BRN-3A) (BRN-3.0).
GN POU4F1 OR BRN3A OR BRN3 OR BRN-3.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.

RC MEDLINE=94215319; PubMed=8162704;
RA Theil T., Zechner U., Rieth C., Adolph S., Moeroy T.;
RT "Chromosomal localization and sequences of the murine Brn-3 family of
RT developmental control genes.";
RL Cytogenet. Cell Genet. 66:267-271(1994).

RL Nucleic Acids Res. 18:1634-1634(1990).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR WHICH MAY PLAY A ROLE IN
CC THE REGULATION OF SPECIFIC GENE EXPRESSION WITHIN A SUBSET OF
CC NEURONAL LINEAGES. MAY PLAY A ROLE IN DETERMINING OR MAINTAINING
CC THE IDENTITIES OF A SMALL SUBSET OF VISUAL SYSTEM NEURONS.

CC -1- TISSUE SPECIFICITY: BRAIN, PERIPHERAL SENSORY NERVOUS SYSTEM AND
CC RETINA. IN THE ADULT NERVOUS SYSTEM BRN-3.0 PREDOMINATES IN THE
CC MEDIAL HABENULA, SUPERFICIAL GRAY OF THE SUPERIOR COLICULUS, RED
CC NUCLEUS, MESENCEPHALIC NUCLEUS OF THE TRIGEMINAL GANGLION, NUCLEUS
CC AMBIGUUS, INFERIOR OLIVARY NUCLEUS, AND PERIPHERAL SENSORY
CC GANGLIA.

CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS
CC TO CLASS-4 POU.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: S69350; AAB30577.2; -
DR EMBL: X51959; CAA36218.1; -
DR PIR: S09237; S09237.
DR HSSP: P10037; 1A07.

DR MGD: MGI:102525; Pou4f1.
DR InterPro: IPR000327; -
DR InterPro: IPR001356; -
DR Pfam: PF00046; homeobox.1.

DR Pfam: PF00157; pou.1
DR PRINTS: PR00028; POU4F1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00035; POU_1; 1.
DR PROSITE: PS00465; POU_2; 1.

DR PROSITE: PS50071; HOMEBOX_2; 1.
CC Transcription regulation; Nuclear protein; DNA-binding; Homeobox;
KW Developmental protein.

FT DOMAIN 57 66 POU-TV BOX.
FT DOMAIN 100 108 POLY-HIS.
FT DOMAIN 130 187 GLY-RICH.
FT DOMAIN 205 258 ALA-RICH.

FT DOMAIN 265 339 POU.
FT DNA_BIND 357 416 HOMEBOX.
SO SEQUENCE 421 AA: 42781 MW: 34EC99D789EBE939 CRC64;

```

FT TRANSMEM 114 133 POTENTIAL.
FT TRANSMEM 244 265 POTENTIAL.
FT TRANSMEM 279 295 POTENTIAL.
FT DOMAIN 140 148 HIS-RICH; COULD BE INVOLVED IN
FT DOMAIN 163 169 COORDINATION OF COBLT IONS.
FT DOMAIN 163 169 HIS-RICH; COULD BE INVOLVED IN
FT CONFLICT 227 227 G -> E (IN REF. 1).
FT CONFLICT 333 334 HI -> V (IN REF. 1).
FT CONFLICT 424 424 I -> V (IN REF. 1).
SQ SEQUENCE 439 AA: 48154 MW: AC88AAASFE2EAAED CRC64:

```

```

Query Match
Best Local Similarity 58.3%; Score 48; DB 1; Length 439;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy 1 HGHEDQHGSHG 12
Db 134 HNDDEHGHGHG 145

```

```

RESULT 12
KE4L_CAEEL STANDARD; PRT; 515 AA.
AC 09XT07;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL KE4-LIKE PROTEIN H13N06.5 IN CHROMOSOME X.
GN H13N06.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69942; CAB17070.1; -
DR WormRep; H13N06.5; CE18815.
KW Hypothetical protein; Transmembrane; Glycoprotein.
FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 49 69 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 366 406 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 463 483 POTENTIAL.
FT DOMAIN 92 182 HIS-RICH.
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 515 AA: 55500 MW: 17D7B54FAE1DAAF CRC64:

```

```

Query Match
Best Local Similarity 58.3%; Score 47; DB 1; Length 515;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Oy 1 HGHEDQHGSHG 12
Db 275 HGHSHSHGHG 286

```

```

RESULT 13
KE4_HUMAN STANDARD; PRT; 469 AA.
ID KE4_HUMAN
AC 092504; Q9UI00;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HISTIDINE-RICH MEMBRANE PROTEIN KE4.
GN HKE4 OR RING5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97001166; PubMed=8812499;
RA Ando A., Kikuli Y.Y., Shigenari A., Kawata H., Okamoto N., Shilina T.,
RA Chen L., Ikemura T., Abe K., Kimura M., Inoko H.;
RT "CDNA cloning of the human homologues of the mouse Ke4 and Ke6 genes
RT at the centromeric end of the human MHC region.";
RL Genomics 35:600-602(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Vergara A., Lana I., Corella A., de Miguel C., Migliaccio M.,
RA Encio I.;
RT "Molecular cloning and characterization of the human KE4 gene and 5'
RT flanking region.";
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Tubby B.;
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- TISSUE SPECIFICITY: MAJOR EXPRESSION IN PLACENTA, LONG, KIDNEY
CC AND PANCREAS.
CC -----
CC -1- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D82060; BA011528.1; -
DR EMBL: AF117221; AAD12305.1; -
DR EMBL: AL031228; CAA20238.1; -
DR MIM: 601416; -
KW Transmembrane; Glycoprotein.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 138 158 POTENTIAL.
FT TRANSMEM 169 189 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.
FT TRANSMEM 417 436 POTENTIAL.
FT DOMAIN 30 114 HIS-RICH.
FT DOMAIN 238 263 HIS-RICH.
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 280 280 A -> G (IN REF. 1 AND 2).
FT CONFLICT 280 280 E -> T (IN REF. 1 AND 2).
FT CONFLICT 376 376 S -> T (IN REF. 1 AND 2).
FT CONFLICT 397 469 CALTEGAGVSGEIVAGAGPGWVLPFTAGGFTYATVSVLP
FT ELTRASPLOSLEVLGIVGVIMVLIATHE -> VEFSL
FT KEQWTVKIOVVYVLAGSCHLLQVALST (IN REF. 1
AND 2).

```

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003451: AAF46615.1: -.
DR FlyBase: FBgn0030227; CG9732.
DR InterPro: IPR002395; -.
DR PRINTS: PR00334; KININOGEN.
SQ SEQUENCE 989 AA; 98851 MM; 642726D9EDADECB4 CRC64;

Query Match 68.0%; Score 51; DB 5; Length 989;
Best Local Similarity 66.7%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HGHGEOHGHG 12
DB 171 HGHGSHGNGHG 182

Search completed: July 6, 2001, 09:25:52
Job time: 988 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:16 : Search time 113.68 Seconds
(without alignments)
6.399 Million cell updates/sec

Title: US-09-437-912-2
Perfect score: 68
Sequence: 1 LDDDLHGGHGV 12

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.GeneSeq_0601.*
2: /SID88/gcgdata/geneSeq/geneSeqp/AA1980.DAT.*
3: /SID88/gcgdata/geneSeq/geneSeqp/AA1981.DAT.*
4: /SID88/gcgdata/geneSeq/geneSeqp/AA1982.DAT.*
5: /SID88/gcgdata/geneSeq/geneSeqp/AA1983.DAT.*
6: /SID88/gcgdata/geneSeq/geneSeqp/AA1984.DAT.*
7: /SID88/gcgdata/geneSeq/geneSeqp/AA1985.DAT.*
8: /SID88/gcgdata/geneSeq/geneSeqp/AA1986.DAT.*
9: /SID88/gcgdata/geneSeq/geneSeqp/AA1987.DAT.*
10: /SID88/gcgdata/geneSeq/geneSeqp/AA1988.DAT.*
11: /SID88/gcgdata/geneSeq/geneSeqp/AA1989.DAT.*
12: /SID88/gcgdata/geneSeq/geneSeqp/AA1990.DAT.*
13: /SID88/gcgdata/geneSeq/geneSeqp/AA1991.DAT.*
14: /SID88/gcgdata/geneSeq/geneSeqp/AA1992.DAT.*
15: /SID88/gcgdata/geneSeq/geneSeqp/AA1993.DAT.*
16: /SID88/gcgdata/geneSeq/geneSeqp/AA1994.DAT.*
17: /SID88/gcgdata/geneSeq/geneSeqp/AA1995.DAT.*
18: /SID88/gcgdata/geneSeq/geneSeqp/AA1996.DAT.*
19: /SID88/gcgdata/geneSeq/geneSeqp/AA1997.DAT.*
20: /SID88/gcgdata/geneSeq/geneSeqp/AA1998.DAT.*
21: /SID88/gcgdata/geneSeq/geneSeqp/AA1999.DAT.*
22: /SID88/gcgdata/geneSeq/geneSeqp/AA2000.DAT.*
23: /SID88/gcgdata/geneSeq/geneSeqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	12	21	AAV81993
2	68	100.0	20	17	AAW07628
3	68	100.0	28	21	AAV81996
4	68	100.0	47	21	AAV93345
5	68	100.0	55	21	AAV93346
6	68	100.0	62	21	AAV93348
7	68	100.0	63	16	AAV75186
8	68	100.0	83	21	AAV93347
9	68	100.0	94	21	AAV93351
10	68	100.0	131	16	AAV75181
11	68	100.0	179	21	AAV93353

12	68	100.0	186	21	AAV93349	Light chain of hum
13	68	100.0	255	21	AAV93342	Light chain of hum
14	64	94.1	16	21	AAV81998	Human two-chain hi
15	42	61.8	57	21	AAV01773	Human secreted pro
16	42	61.8	107	18	AAW27561	Human cytoctatin I
17	42	61.8	107	18	AAW26581	Human cytoctatin I
18	42	61.8	107	20	AAV49535	Human cytoctatin I
19	42	61.8	107	20	AAW83929	Human growth inhib
20	42	61.8	135	18	AAW30891	Human cytoctatin I
21	42	61.8	135	20	AAV32504	Human cytoctatin I
22	42	61.8	135	21	AAV92910	Human retinoid bin
23	42	61.8	135	22	AAW60659	Human cellular ret
24	41	60.3	538	12	AAV15502	Cytochrome P 450 C
25	40	58.8	217	21	AAV28213	Arabidopsis thalia
26	40	58.8	278	21	AAV28212	Arabidopsis thalia
27	40	58.8	413	21	AAV41846	Arabidopsis thalia
28	40	58.8	1025	16	AAV70126	Human ORFX ORF1610
29	39	57.4	540	21	AAV90601	Serum opacity fact
30	39	57.4	540	21	AAV90602	Candida tropicalis
31	39	57.4	613	20	AAV93823	E. coli GUS protei
32	39	57.4	613	20	AAV93828	Human GUS protei
33	39	57.4	613	21	AAV28407	Escherichia coli b
34	39	57.4	909	15	AAV50092	Humanised anti-CEA
35	38	55.9	238	22	AAV61611	Human protein HP03
36	38	55.9	248	21	AAV61611	Human ORFX ORF1437
37	38	55.9	593	19	AAV62835	Zea mays anti-micro
38	38	55.9	688	20	AAV00241	Enterococcus faeca
39	38	55.9	2032	20	AAV00238	Enterococcus faeca
40	38	55.9	2032	20	AAV00240	Enterococcus faeca
41	38	55.9	2032	20	AAV00242	Enterococcus faeca
42	38	55.9	142	21	AAV1594	Arabidopsis thalia
43	37	54.4	148	21	AAV1593	Arabidopsis thalia
44	37	54.4	183	21	AAV1592	Arabidopsis thalia
45	37	54.4	183	21	AAV1592	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAV81993	standard; peptide: 12 AA.
ID	AAV81993
AC	AAV81993;
XX	
DT	16-OCT-2000 (first entry)
DE	Human high molecular weight kininogen domain 5 fragment #2.
XX	
KW	Human: high molecular weight kininogen; HK;
KW	two-chain high molecular weight kininogen; Hka;
KW	angiogenesis inhibition; tumour; cancer; ocular disorder;
KW	rheumatoid arthritis; endothelial cell apoptosis.
OS	Homo sapiens.
XX	
PN	WO200027866-A1.
PD	18-MAY-2000.
XX	
PF	05-NOV-1999; 99WO-US26419.
XX	
PR	10-NOV-1998; 98US-0107833.
XX	
PA	(UTEM) UNIV TEMPLE.
PA	(MCCR/) MCCRAE R K.
XX	
PI	MCCRAE RK;
XX	
DR	WPI. 2000-376483/32.
XX	
PT	A pharmaceutical composition used to inhibit angiogenesis, inhibit endothelial cell proliferation, and induce endothelial cell apoptosis

PT -
XX
PS Claim 5; Page 28; 52pp; English.
XX

CC The present sequence is derived from human high molecular
CC weight kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds
CC with high affinity to endothelial cells, where it is cleaved to
CC two-chain high molecular weight kininogen (Hka) by plasma kallikrein.
CC Hka or a synthetic compound comprising part or all of the present
CC sequence may be used in a pharmaceutical composition for inhibiting
CC angiogenesis. Angiogenesis occurs in a number of disease states, such
CC as tumour formation and expansion, and certain ocular disorders. It can
CC also occur in a rheumatoid joint, hastening joint destruction by
CC allowing an influx of leukocytes. The composition may inhibit
CC angiogenesis by inhibiting endothelial cell proliferation or by
CC inducing endothelial cell apoptosis. Peptides used in the composition
CC may be recombinant peptides, natural peptides, or synthetic peptides.
CC They may also be chemically synthesised, using, for example, solid
CC phase synthesis methods.
XX
SQ

Sequence 12 AA;

Query Match 100.0%; Score 68; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDDLHHOGGHV 12
1 | | | | | | | | | |
Db 1 ldddlhggghv 12

RESULT 2

AAW07628
ID AAW07628 standard; peptide; 20 AA.

AC AAW07628;

DR 04-FEB-1997 (first entry)

XX Human high polymer quininogen L-chain derived peptide.

XX Human; high polymer; quininogen; L-chain.

OS Homo sapiens.

PN JP08208692-A.

PD 13-AUG-1996.

PF 28-SEP-1995; 95JP-0276418.

PR 28-SEP-1994; 94JP-0259451.

XX (SDMU) SUMITOMO SEIYAKU KK.

DR WPI; 1996-421988/42.

PT Cell adhesion inhibiting peptide(s), used as cancer metastasis
PT inhibitor - comprises partial amino acid sequence of human high
PT polymer quininogen L chain

PS Example: Page 8; 14pp; Japanese.

CC The present peptide is derived from residues 402-498 of the human
CC high polymer quininogen L-chain. It was synthesised using a solid
CC phase method, and purified using a YMC-DO5-120A-S15/13 column.

XX Sequence 20 AA;

Query Match 100.0%; Score 68; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDDLHHOGGHV 12
1 | | | | | | | | | |
Db 3 ldddlhggghv 14

RESULT 3

AAW81996
ID AAW81996 standard; peptide; 28 AA.

AC AAW81996;

DR 16-OCT-2000 (first entry)

XX Human high molecular weight kininogen domain 5 fragment #5.

XX Human; high molecular weight kininogen; HK;

KW two-chain high molecular weight kininogen; Hka;

KW angiogenesis inhibition; tumour; cancer; ocular disorder;

XX rheumatoid arthritis; endothelial cell apoptosis.

OS Homo sapiens.

PN WO200027866-A1.

PD 18-MAY-2000.

PF 05-NOV-1999; 99WO-US26419.

PR 10-NOV-1998; 98US-0107833.

XX (UTEM) UNIV TEMPLE.

PA (MCCR/) MCCRAE R K.

PI McCrae RK;

DR WPI; 2000-376483/32.

XX A pharmaceutical composition used to inhibit angiogenesis, inhibit

PT endothelial cell proliferation, and induce endothelial cell apoptosis

PT -

PS Claim 8; Page 28; 52pp; English.

CC The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain
CC high molecular weight kininogen (Hka) by plasma kallikrein. Hka or a
CC synthetic compound comprising the present sequence may be used in a
CC pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
CC occurs in a number of disease states, such as tumour formation and
CC expansion, and certain ocular disorders. It can also occur in a
CC rheumatoid joint, hastening joint destruction by allowing
CC an influx of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial
CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.

XX Sequence 28 AA;

Query Match 100.0%; Score 68; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDDLHHOGGHV 12
1 | | | | | | | | | |
Db 17 ldddlhggghv 28

	RESULT	4	
AAV93345	ID	AAV93345 standard; peptide: 47 AA.	
XX	AC	AAV93345;	
DT	04-SEP-2000	(first entry)	
DE	Light chain of human high molecular weight kininogen fragment.		
KW	Human; high molecular weight kininogen; glycoprotein; endothelial cell; plasma kallikrein; heavy chain; light chain; analogue; angiogenesis; endothelial cell proliferation; endothelial cell migration; vitronectin. 		
OS	Synthetic.		
OS	Homo sapiens.		
PN	WO200027415-A2.		
PD	18-MAY-2000.		
PE	09-NOV-1999; 99WO-US26377.		
PR	10-NOV-1998; 98US-0107844.		
PA	(UTEM) UNIV TEMPLE. (DUPO) DUPONT PHARM CO. (COLM/) COLMAN W R. (MOUSA/) MOUSA A S.		
PI	Coleman WR, Mousa AS;		
DR	WPI: 2000-376306/32.		
PT	Method for inhibiting endothelial cell proliferation, using compound that inhibit endothelial cell migration - 		
PS	Claim 3; Page 36; 41pp: English.		
CC	The present sequence represents a fragment of the light chain of human high molecular weight kininogen. It is used to produce compounds of the invention. High molecular weight kininogen is a 120 KDa glycoprotein which binds with high affinity to endothelial cells, where it is cleaved by plasma kallikrein into heavy and light chains. Analogues of high molecular weight kininogen are used in the method of the invention. The specification describes a method of inhibiting endothelial cell proliferation. The method comprises contacting endothelial cells with a compound containing high molecular weight kininogen analogues. The method and the compounds can also be used for inhibiting endothelial cell proliferation. The compounds can also be used for inhibiting angiogenesis. The compounds can also be used to inhibit migration of endothelial cells to vitronectin.		
Sequence	47 AA:		
OY	Query Match	100.0%; Score 68; DB 21; Length 47; Best Local Similarity 100.0%; Pred. NO. 4,7e-05;	
Matches	12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
DQ	1 IDDDLHOGGHV 12 		
DB	6 ldddlehgghv 17		
Result	5		
ID	AAV93346		
XX	AAV93346 standard; peptide: 55 AA.		
AC	AAV93346;		
DT	04-SEP-2000 (first entry)		

DE	Light chain of human high molecular weight kininogen analogue.
KW	Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KM	plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KV	endothelial cell proliferation; endothelial cell migration; vitronectin.
XX	
OS	Synthetic.
OS	Homo sapiens.
PN	WO200027415-A2.
PD	18-MAY-2000.
PF	09-NOV-1999; 99WO-US26377.
PR	10-NOV-1998; 98US-0107844.
PA	(UTEM) UNIV TEMPLE.
PA	(DUPO) DUPONT PHARM CO.
PA	(COLM/) COLMAN W R.
PA	(MOUS/) MOUSA A S.
PI	Colman WR, Mousa AS;
DR	WPI; 2000-376306/32.
PT	Method for inhibiting endothelial cell proliferation, using compound
PT	that inhibit endothelial cell migration -
PS	Claim 4; Page 36; 41pp; English.
CC	The present sequence represents an analogue of the light chain of human
CC	high molecular weight kininogen. High molecular weight kininogen is a
CC	120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC	where it is cleaved by plasma kallikrein into heavy and light chains.
CC	Analogues of high molecular weight kininogen are used in the method
CC	of the invention. The specification describes a method of inhibiting
CC	endothelial cell proliferation. The method comprises contacting
CC	endothelial cells with a compound containing high molecular weight
CC	kininogen analogues. The method and the compounds can be used for
CC	inhibiting endothelial cell proliferation. The compounds can also be
CC	used for inhibiting angiogenesis. The compounds can also be used to
CC	inhibit migration of endothelial cells to vitronectin.
XX	
XX	
SQ	Sequence 55 AA:
Query Match	100.0%; Score 68; DB 21; Length 55;
Best Local Similarity	100.0%; Pred. No. 5.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.	
OY	1 LDDDLHOGGHV 12
DB	42 ldddlehgghv 53
RESULT 6	
AAV93348	
ID	AAV93348 standard; peptide; 62 AA.
AC	AAV93348;
XX	
DT	04-SEP-2000 (first entry)
XX	
XX	
XX	Light chain of human high molecular weight kininogen analogue.
KW	Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KM	plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KV	endothelial cell proliferation; endothelial cell migration; vitronectin.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	

PN WO200027415-A2.
XX
XX 18-MAY-2000.
XX
XX 09-NOV-1999; 99WO-US26377.
XX
XX 10-NOV-1998; 98US-0107844.
XX
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
PI
XX
XX WPI: 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX
XX
XX Claim 6; Page 37; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
XX Sequence 62 AA:
SQ
Query Match 100.0%; Score 68; DB 21; Length 62;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LDDDEHOGGHV 12
DB 21 ldddlehgghv 32
RESULT 7
AA75186
ID AA75186 standard; peptide: 63 AA.
XX
XX AA75186;
AC
XX
XX 05-DEC-1995 (first entry)
DT
XX
XX Partial peptide of human HMW kininogen fragment 2.
DE
XX
XX high molecular weight; kininogen; fragment; 1,2; 1; 2; partial;
KW wound treating agent; bovine; growth promotion; fibroblast.
XX
XX Homo sapiens.
OS
XX
XX JP07082172-A.
PN
XX
XX 28-MAR-1995.
PD
XX
XX 17-SEP-1993; 93JP-0230616.
PE
XX
XX 17-SEP-1993; 93JP-0230616.
PR
XX
XX (FARH) HOECHST JAPAN KK.
PA
XX
XX WPI: 1995-158909/21.
DR

XX
XX A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
XX
XX Claim 8; Page 8; 8pp; Japanese.
XX
XX AA75186 is a partial peptide corresponding to human kininogen
CC fragment 1, amino acids 458-520. Partial peptides of bovine and
CC human kininogen fragments 1,2, 1 and 2, are used in wound treating
CC agent comps. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
XX
XX Sequence 63 AA:
SQ
Query Match 100.0%; Score 68; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 6.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LDDDEHOGGHV 12
DB 22 ldddlehgghv 33
RESULT 8
AA93347
ID AA93347 standard; peptide: 83 AA.
XX
XX AA93347;
AC
XX
XX 04-SEP-2000 (first entry)
DT
XX
XX Light chain of human high molecular weight kininogen analogue.
DE
XX
XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
XX
XX Synthetic.
OS
XX
XX Homo sapiens.
OS
XX
XX WO200027415-A2.
PN
XX
XX 18-MAY-2000.
PD
XX
XX 09-NOV-1999; 99WO-US26377.
PE
XX
XX 10-NOV-1998; 98US-0107844.
PR
XX
XX (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOUS/) MOUSA A S.
XX
XX Colman WR, Mousa AS;
PI
XX
XX WPI: 2000-376306/32.
XX
XX Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX
XX
XX Claim 5; Page 37; 41pp; English.
XX
XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for

CC Inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 83 AA;

Query Match 100.0%; Score 68; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 8.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDLEHOGGHV 12
|||
DB 42 ldddlehgghv 53

RESULT 9

AAV93351
ID AAV93351 standard; peptide; 94 AA.

AC AAV93351;

DT 04-SEP-2000 (first entry)

XX Light chain of human high molecular weight kininogen analogue.

XX Human: high molecular weight kininogen; glycoprotein; endothelial cell;

KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

OS Homo sapiens.

PN WO200027415-A2.

XX 18-MAY-2000.

XX 09-NOV-1999; 99WO-US26377.

XX 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.

PA (DUPO) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

PI Colman WR, Mousa AS;

DR WPI; 2000-376306/32.

XX Method for inhibiting endothelial cell proliferation, using compound

PT that inhibit endothelial cell migration -

XX Claim 8; Page 39; 41pp; English.

XX The present sequence represents an analogue of the light chain of human

CC high molecular weight kininogen. High molecular weight kininogen is a

CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,

CC where it is cleaved by plasma kallikrein into heavy and light chains.

CC Analogues of high molecular weight kininogen are used in the method

CC of the invention. The specification describes a method of inhibiting

CC endothelial cell proliferation. The method comprises contacting

CC endothelial cells with a compound containing high molecular weight

CC kininogen analogues. The method and the compounds can be used for

CC inhibiting endothelial cell proliferation. The compounds can also be

CC used for inhibiting angiogenesis. The compounds can also be used to

CC inhibit migration of endothelial cells to vitronectin.

XX Sequence 94 AA;

XX Query Match 100.0%; Score 68; DB 21; Length 94;

XX Best Local Similarity 100.0%; Pred. No. 0.0001;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LDDLEHOGGHV 12
|||
DB 42 ldddlehgghv 53

RESULT 10

AAV75181

ID AAV75181 standard; peptide; 131 AA.

AC AAV75181;

DT 05-DEC-1995 (first entry)

XX Partial peptide of human HMW kininogen fragment 1.2.

XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;

KW wound treating agent; human; growth promotion; fibroblast.

XX Homo sapiens.

OS JP07082172-A.

XX 28-MAR-1995.

XX 17-SEP-1993; 93JP-0230616.

XX 17-SEP-1993; 93JP-0230616.

XX (FARH) HOECHST JAPAN KK.

XX WPI; 1995-158909/21.

XX A wound treating agent contg. a partial peptide of kininogen -

PT have growth promotion activity of fibroblasts.

XX Claim 7; Page 7; 8pp; Japanese.

XX AAR75181 is a partial peptide corresponding to human kininogen

CC fragment 1.2, amino acids 390-520. Partial peptides of bovine and

CC human kininogen fragments 1.2, 1 and 2, are used in wound treating

CC agent compsns. and act as the active component. The fragments are

CC useful in wound treating because they have growth promotion activity

CC on fibroblasts.

XX Sequence 131 AA;

XX Query Match 100.0%; Score 68; DB 16; Length 131;

XX Best Local Similarity 100.0%; Pred. No. 0.00015;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDLEHOGGHV 12
|||
DB 90 ldddlehgghv 101

XX RESULT 11

AAV93353

ID AAV93353 standard; peptide; 179 AA.

AC AAV93353;

DT 04-SEP-2000 (first entry)

XX Light chain of human high molecular weight kininogen analogue.

XX Human: high molecular weight kininogen; glycoprotein; endothelial cell;

KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

KW endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

Query Match 61.8%; Score 42; DB 21; Length 57;
 Best Local Similarity 54.5%; Pred. No. 1.9;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 DDDLEHOGGHV 12
 |::||| |
 Db 40 dkelehqghm 50

Search completed: July 6, 2001, 09:09:17
 Job time: 123 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:20 ; Search time 56.74 Seconds
(without alignments)
4.260 Million cell updates/sec

Title: US-09-437-912-2

Perfect score: 68

Sequence: 1 LDDLEHOGSHV 12

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 segs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	61.8	106	2	US-08-820-825-14
2	42	61.8	106	4	US-09-307-817-14
3	42	61.8	107	1	US-08-409-731A-2
4	42	61.8	107	2	US-08-470-298B-2
5	42	61.8	107	2	US-09-023-073A-2
6	42	61.8	135	2	US-08-820-825-2
7	42	61.8	135	3	US-08-899-031-1
8	42	61.8	135	4	US-09-307-817-2
9	36	52.9	604	2	US-08-468-576B-12
10	36	52.9	604	2	US-08-468-579B-12
11	36	52.9	604	3	US-08-468-577B-12
12	36	50.0	447	1	US-08-476-008-67
13	34	50.0	447	1	US-08-306-063-67
14	34	50.0	447	1	US-08-833-485-67
15	34	50.0	560	1	US-08-647-484-2
16	34	50.0	560	1	US-08-647-481-2
17	34	50.0	560	1	US-08-430-033A-2
18	34	50.0	560	2	US-08-805-118-4
19	34	50.0	560	5	PCT-US96-05792-2
20	34	50.0	600	6	5268463-2
21	34	50.0	602	6	US-08-882-704A-5
22	34	50.0	602	6	5432081-2
23	34	50.0	832	3	US-08-630-820-7
24	34	50.0	1079	3	US-09-136-652-2
25	34	50.0	1110	1	US-08-118-441-29
26	34	50.0	1110	3	US-08-338-579A-29
27	34	50.0	1110	5	PCT-US94-09851-29

28	34	50.0	1408	1	US-08-612-521-2	Sequence 2, Appl1
29	33	48.5	26	2	US-08-943-583-6	Sequence 6, Appl1
30	33	48.5	27	2	US-08-943-583-1	Sequence 1, Appl1
31	33	48.5	78	1	US-07-929-206-4	Sequence 4, Appl1
32	33	48.5	78	2	US-08-313-185-44	Sequence 44, Appl1
33	33	48.5	78	2	US-08-459-499-4	Sequence 4, Appl1
34	33	48.5	78	4	US-09-082-614A-44	Sequence 14, Appl1
35	33	48.5	370	4	US-08-878-989-19	Sequence 44, Appl1
36	33	48.5	370	4	US-09-272-796-19	Sequence 19, Appl1
37	33	48.5	455	2	US-08-272-255-14	Sequence 14, Appl1
38	33	48.5	455	2	PCT-US95-08365-14	Sequence 14, Appl1
39	33	48.5	472	4	US-08-895-590-11	Sequence 11, Appl1
40	33	48.5	500	2	US-08-987-519-11	Sequence 1, Appl1
41	33	48.5	652	2	US-08-313-185-53	Sequence 17, Appl1
42	33	48.5	652	2	US-08-459-499-17	Sequence 53, Appl1
43	33	48.5	652	4	US-09-082-614A-53	Sequence 53, Appl1
44	33	48.5	726	2	US-08-313-185-49	Sequence 49, Appl1
45	33	48.5	726	2	US-08-459-499-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-08-820-825-14
Sequence 14, Application US/08820825
Patent No. 5945309
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER L.
APPLICANT: DILLON, PATRICK
TITLE OF INVENTION: CYTOSTATIN III
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/820, 825
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PE222
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-820-825-14

Query Match 61.8%; Score 42; DB 2; Length 106;
Best Local Similarity 54.5%; Pred. No. 2.3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGSHV 12

I : : : : : I :

DB 39 DKEIHQGNHM 49

RESULT 2

US-09-307-817-14

Sequence 14, Application US/09307817
Patent No. 6232291

GENERAL INFORMATION:

APPLICANT: NI, JIAN

APPLICANT: YU, GUO-LIANG

APPLICANT: GENTZ, REINER L.

TITLE OF INVENTION: CYTOSTATIN III

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: HUMAN GENOME SCIENCES, INC.

STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/307,817

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/820,825

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: BROOKES, ANDERS A.

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PF222

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-307-817-14

Query Match

Best Local Similarity 54.5%; Pred. No. 2.3;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGSHV 12

DB 39 DKEIHQGNHM 49

RESULT 3

US-08-409-731A-2

Sequence 2, Application US/08409731A

Patent No. 5658758

GENERAL INFORMATION:

APPLICANT: NI, JIAN

APPLICANT: YU, GUO-LIANG

APPLICANT: GENTZ, REINER

APPLICANT: ROSEN, CRAIG A.

TITLE OF INVENTION: CYTOSTATIN I

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: HUMAN GENOME SCIENCES, INC.

STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/409,731A

FILING DATE: 24-MAR-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Robert H

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PF175

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8512

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-409-731A-2

Query Match

Best Local Similarity 54.5%; Pred. No. 2.3;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGSHV 12

DB 40 DKEIHQGNHM 50

RESULT 4

US-08-470-298B-2

Sequence 2, Application US/08470298B

Patent No. 5844081

GENERAL INFORMATION:

APPLICANT: NI, JIAN

APPLICANT: GENTZ, REINER

APPLICANT: YU, GUO-LIANG

APPLICANT: ROSEN, CRAIG A.

TITLE OF INVENTION: CYTOSTATIN I

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: HUMAN GENOME SCIENCES, INC.

STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE

STATE: MD

COUNTRY: US

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,298B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BROOKES, ALLAN A.

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PF175D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8512

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-470-298B-2

Query Match 61.8%; Score 42; DB 2; Length 107;
Best Local Similarity 54.5%; Pred. No. 2.3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGGHV 12
| : : | | | | :
DB 40 DKEIHOGNHM 50

RESULT 5

US-09-023-073A-2
Sequence 2, Application US/09023073A
Patent No. 5977309

GENERAL INFORMATION:
APPLICANT: NI, Jian

APPLICANT: Gentz, Reiner

APPLICANT: Yu, Guo-Liang

APPLICANT: Rosen, Craig A

TITLE OF INVENTION: Cytostatin I

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.

STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE

STATE: MARYLAND

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023.073A

FILING DATE: 13-FEB-1998

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Wales, Michele M.

REGISTRATION NUMBER: P-43,975

REFERENCE/DOCKET NUMBER: PF17502

TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5772

TELEPHONE: 301-309-8439

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-023-073A-2

Query Match 61.8%; Score 42; DB 2; Length 107;
Best Local Similarity 54.5%; Pred. No. 2.3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGGHV 12
| : : | | | | :
DB 40 DKEIHOGNHM 50

RESULT 6

US-08-820-825-2
Sequence 2, Application US/08820825
Patent No. 5945309

GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER L.
APPLICANT: DILLON, PATRICK
TITLE OF INVENTION: CYTOSTATIN III
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.

STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/820,825

FILING DATE: 19-MAR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PF222

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512

TELEPHONE: (301) 309-8504

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 135 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-820-825-2

Query Match 61.8%; Score 42; DB 2; Length 135;
Best Local Similarity 54.5%; Pred. No. 3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGGHV 12
| : : | | | | :
DB 40 DKEIHOGNHM 50

RESULT 7

US-08-899-031-1
Sequence 1, Application US/08899031
Patent No. 6046027

GENERAL INFORMATION:
APPLICANT: Bandman, Olga

APPLICANT: Guegler, Karl J.

APPLICANT: Shah, Purvu

TITLE OF INVENTION: HUMAN RETINOID BINDING PROTEIN

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,031

FILING DATE: Herewith

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0349 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAT01
CLONE: 367304
US-08-899-031-1

Query Match 61.8%; Score 42; DB 3; Length 135;
Best Local Similarity 54.5%; Pred. No. 3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGHV 12
1 : : : : :
DB 40 DKEIHHGNHM 50

RESULT 8

US-09-307-817-2
Sequence 2, Application US/09307817
Patent No. 6232291
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: YU, GUO-LIANG
APPLICANT: GENTZ, REINER L.
TITLE OF INVENTION: CYTOSTATIN III
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/307,817
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/820,825
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF222
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-307-817-2

Query Match 61.8%; Score 42; DB 4; Length 135;
Best Local Similarity 54.5%; Pred. No. 3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGHV 12
1 : : : : :
DB 40 DKEIHHGNHM 50

RESULT 9

US-08-468-576B-12
Sequence 12, Application US/08468576B
Patent No. 595345
GENERAL INFORMATION:
APPLICANT: Rabio, Daniel
TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: Sprung Kramer Schaefer & Briscoe
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System 7.5
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,576B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,276
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,646
FILING DATE: 08-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/715,181
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/441,703
FILING DATE: 04-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/312,543
FILING DATE: 17-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: MDI 251.7-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 604 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-576B-12

Query Match 52.9%; Score 36; DB 2; Length 604;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 DDLEHOGGH 11
1 1 1 1 1
DB 351 DFEHGGH 359

RESULT 10
US-08-468-579B-12

; Sequence 12, Application US/08468579B

; Patent No. 5981700

; GENERAL INFORMATION:

; APPLICANT: Rabin, Daniel

; TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS

; TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sprung Kramer Schaefer & Briscoe

; STREET: 660 White Plains Road

; CITY: Tarrytown

; STATE: New York

; COUNTRY: USA

; ZIP: 10591-5144

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: System 7.5

; SOFTWARE: WordPerfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468, 579B

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/239, 276

; FILING DATE: 05-MAY-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/872, 646

; FILING DATE: 08-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/715, 181

; FILING DATE: 14-JUN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/441, 703

; FILING DATE: 04-DEC-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/312, 543

; FILING DATE: 17-FEB-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Kurt G. Briscoe

; REGISTRATION NUMBER: 33,141

; REFERENCE/DOCKET NUMBER: MDI 251.5-KGB

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (914) 332-1700

; TELEFAX: (914) 332-1844

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 604 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-08-468-579B-12

OY 3 DDLEHOGGH 11
1 1 1 1 1
DB 351 DFEHGGH 359

Query Match 52.9%; Score 36; DB 2; Length 604;

Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 DDLEHOGGH 11
1 1 1 1 1
DB 351 DFEHGGH 359

RESULT 11

US-08-468-577B-12

; Sequence 12, Application US/08468577B

; Patent No. 6001804

; GENERAL INFORMATION:

; APPLICANT: Rabin, Daniel

; TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS

; TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sprung Kramer Schaefer & Briscoe

; STREET: 660 White Plains Road

; CITY: Tarrytown

; STATE: New York

; COUNTRY: USA

; ZIP: 10591-5144

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: System 7.5

; SOFTWARE: WordPerfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/468, 577B

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/239, 276

; FILING DATE: 05-MAY-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/872, 646

; FILING DATE: 08-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/715, 181

; FILING DATE: 14-JUN-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/441, 703

; FILING DATE: 04-DEC-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/312, 543

; FILING DATE: 17-FEB-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Kurt G. Briscoe

; REGISTRATION NUMBER: 33,141

; REFERENCE/DOCKET NUMBER: MDI 251.8-KGB

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (914) 332-1700

; TELEFAX: (914) 332-1844

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 604 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-08-468-577B-12

OY 3 DDLEHOGGH 11
1 1 1 1 1
DB 351 DFEHGGH 359

Query Match 52.9%; Score 36; DB 3; Length 604;

Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 DDLEHOGGH 11
1 1 1 1 1
DB 351 DFEHGGH 359

RESULT 12

US-08-476-008-67

; Sequence 67, Application US/08476008

; Patent No. 5627061

; GENERAL INFORMATION:

; APPLICANT: Barry, Gerard F.

; APPLICANT: Kishore, Ganesh M.

; APPLICANT: Padgett, Stephen R.

; APPLICANT: Stallings, William C.

; TITLE OF INVENTION: Glycosyltransferase

; NUMBER OF SEQUENCES: 69

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
;; STREET: 700 Chesterfield Village Parkway
;; CITY: St. Louis
;; STATE: Missouri
;; COUNTRY: USA
;; ZIP: 63198
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/476,008
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/306,063
;; FILING DATE: 13-SEP-1994
;; APPLICATION NUMBER: US 07/749,611
;; FILING DATE: 28-AUG-1991
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 07/576,537
;; FILING DATE: 31-AUG-1990
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hoerner Jr., Dennis R.
;; REGISTRATION NUMBER: 30,914
;; REFERENCE/DOCKET NUMBER: 38-21(10660)A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (314)537-6099
;; TELEFAX: (314)537-6047
;; INFORMATION FOR SEQ ID NO: 67:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 447 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-476-008-67

Query Match 50.0%; Score 34; DB 1; Length 447;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGG 10
DB 383 DDGLEIQQG 391

RESULT 13
US-08-306-063-67
; Sequence 67, Application US/08306063
; Patent No. 5633435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyophosate tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/306,063
;; FILING DATE: 13-SEP-1994
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 07/749,611
;; FILING DATE: 28-AUG-1991
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 07/576,537
;; FILING DATE: 31-AUG-1990
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hoerner Jr., Dennis R.
;; REGISTRATION NUMBER: 30,914
;; REFERENCE/DOCKET NUMBER: 38-21(10660)A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (314)537-6099
;; TELEFAX: (314)537-6047
;; INFORMATION FOR SEQ ID NO: 67:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 447 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-306-063-67

Query Match 50.0%; Score 34; DB 1; Length 447;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGG 10
DB 383 DDGLEIQQG 391

RESULT 14
US-08-833-485-67
; Sequence 67, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyophosate tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,485
; FILING DATE: 07-APR-1997
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611

FILED DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-485-67

Query Match 50.0%; Score 34; DB 1; Length 447;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 DDDLEHOGG 10
DB 383 DDLEIQQG 391

RESULT 15
US-08-647-484-2
Sequence 2, Application US/08647484
Patent No. 5618677
GENERAL INFORMATION:
APPLICANT: NI, Binhui
APPLICANT: Paul, Steven M.
TITLE OF INVENTION: HUMAN BRAIN SODIUM DEPENDENT INORGANIC
TITLE OF INVENTION: PHOSPHATE COTRANSPORTER AND RELATED NUCLEIC ACID COMPOUNDS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/647,484
FILING DATE: 14-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,033
FILING DATE: 27-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-10006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-647-484-2
Query Match 50.0%; Score 34; DB 1; Length 560;
Best Local Similarity 41.7%; Pred. No. 3.2e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 LDDLEHOGGHV 12
DB 91 VNNSTTHRGHV 102

Search completed: July 6, 2001, 09:10:20
Job time: 186 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:17:57 : Search time 73.59 seconds
(without alignments)
12.421 Million cell updates/sec

Title: US-09-437-912-2
Perfect score: 68
Sequence: 1 LDDLEHOGGHV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_68:*

1: pir1:*\n2: pir2:*\n3: pir3:*\n4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	644	1 KGHU1	kininogen, HMW pre
2	44	64.7	541	2 T34850	probable acid-CoA
3	43	63.2	174	2 G83712	hypothetical prote
4	42	61.8	683	2 S01433	repressor protein
5	42	61.8	3282	2 E82750	hemagglutinin-like
6	42	61.8	3442	2 E82589	hemagglutinin-like
7	42	61.8	3455	2 B82519	hemagglutinin-like
8	41	60.3	538	1 OACR44	cytochrome P450 52
9	40.5	59.6	536	2 T27668	hypothetical prote
10	40	58.8	1025	2 S69790	fibronectin-bindin
11	40	58.8	1733	2 S27939	tensin - chicken
12	40	58.8	1744	2 A54970	tensin - chicken
13	40	58.8	1792	2 A57075	tensin - chicken
14	39	57.4	325	2 F83503	hypothetical prote
15	39	57.4	534	1 A48529	D-lactate dehydrog
16	39	57.4	571	1 DEECDL	hypothetical prote
17	39	57.4	585	1 E85850	unknown protein T1
18	39	57.4	581	2 G96811	beta-glucuronidase
19	39	57.4	651	2 A26581	GTP pyrophosphokin
20	39	57.4	744	1 KTECG	hypothetical prote
21	39	57.4	744	2 E85929	hypothetical prote
22	38	55.9	81	2 C81931	globulin-10 - maize
23	38	55.9	122	2 C53234	globulin-1 - maize
24	38	55.9	236	2 T01662	probable ATP-bind
25	38	55.9	407	2 T02258	vicillin-like stora
26	38	55.9	489	2 T36100	vicillin-like stora
27	38	55.9	540	2 S21825	globulin-1S, GLB1S
28	38	55.9	573	2 A53234	vicilin-like stora
29	38	55.9	582	2 B53234	

30	38	55.9	818	2 A4858	Na+/H+-exchanging
31	38	55.9	2422	2 T18201	fatty-acid synthas
32	38	55.9	5170	2 T15348	hypothetical prote
33	37.5	55.1	511	2 A37803	beta-fructofuranos
34	37	54.4	64	2 C83929	hypothetical prote
35	37	54.4	179	2 E83601	protease P1PI PA
36	37	54.4	183	2 T51572	hypothetical prote
37	37	54.4	191	2 D96701	unknown protein, 9
38	37	54.4	285	2 B83047	hypothetical prote
39	37	54.4	339	2 T06612	hypothetical prote
40	37	54.4	407	2 T39658	probable mitochond
41	37	54.4	420	2 T39712	hypothetical prote
42	37	54.4	426	2 H64878	probable oxidoredu
43	37	54.4	426	2 C85755	probable oxidoredu
44	37	54.4	427	2 T42516	hypothetical prote
45	37	54.4	473	2 G83932	two-component sens

ALIGNMENTS

RESULT 1
KGHU1
N:kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-chiol proteinase inhibitor; preprokininogen; prokininogen
N:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence-revision 28-May-1986 #text-change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030;
R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A:Title: Isolation of a human cDNA for alpha-2-chiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:85122621
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:K02566; NID:G177889
R:Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and 1
A:Reference number: A92544; MUID:85234582
A:Accession: A25276
A:Molecule type: mRNA
A:Residues: 1-592, 'I', 594-644 <TRK>
A:Cross-references: GB:M1437; NID:G186751; PID:AB59550.1; PID:G386852
R:Auerwald, E.A.; Roessler, D.; Mentel, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A:Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:93223854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 'ANSM', 253-377 <AUE>
A:Note: differences are due to known cloning artifacts
R:Loetsch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kin
A:Reference number: A91153; MUID:86030270
A:Accession: A91153
A:Molecule type: protein
A:Residues: 379-644 <LOT>
A:Note: the bradykinin sequence preceding the light chain sequence was not determined
R:Kellermann, J.; Loetsch, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A:Title: Completion of the primary structure of human high-molecular-mass kininogen.
A:Reference number: A24871; MUID:86108361
A:Accession: A24871
A:Molecule type: protein
A:Residues: 'Z', 20-380 <KEI>
R:Kellermann, J.; Loetsch, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolis, H.S., ed., pp.85-89, Plenum Press, New
A:Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A:Reference number: A27899
A:Accession: A27899

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEI2>
 R:Minidrou, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88209021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A31905; MUID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAE>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plasma
 A:Reference number: A34030; MUID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPis) in inflammatory and
 A:Reference number: S02482; MUID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19; 169-192; 310-314; 381-389 <LENI>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-specific
 A:Reference number: S55239; MUID:95251593
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Stracek, J.; Maechli, F.; Le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; MUID:96033974
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evolution
 A:Reference number: A92545; MUID:85234583
 A:Contents: annotation; gene organization

R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:90255622
 A:Contents: annotation; bradykinin
 A:Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
 C:Comment: Kininogen is a cysteine proteinase inhibitor. takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is 1
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 C:Comment: The residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-References: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379, 390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <RCH>
 F:19-131/Domain: cystatin homology <CV1>
 F:142-253/Domain: cystatin homology <CV2>
 F:264-375/Domain: cystatin homology <CV3>
 F:380-389/Product: lysyl-bradykinin (kallidin I) #status experimental <BDY>
 F:390-644/Product: bradykinin (kallidin II) #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:28-614, 83-94, 107-126, 142-145, 206-218, 229-248, 264-267, 328-340, 351-370/Disulfide bond
 F:48/Binding site: carboxylate (Asp) (covalent) #status absent
 F:169, 205, 294/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 593, 628/Binding site: carboxylate (Thr) (covalent) #status
 F:577/Binding site: carboxylate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 68; DB 1; Length 644;
 Best Local Similarity 100.0%; Pred. No. 0.00058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LDDDEHOGCHV 12
 DB 479 LDDDEHOGCHV 490
 RESULT 2
 T34850
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
 C:Accession: T34850
 R:Olliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
 submitted to the EMBL Data Library, February 1999
 A:Reference number: Z21559
 A:Accession: T34850
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-541
 A:Cross-References: EMBL:AL005478; PIDN:CA836604.1; GSPDB:GN00070; SCOEDB:SC265.17
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC265.17
 C:Superfamily: 4-oxonucleate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: acid-thiol ligase
 F:70-533/Domain: acetate--CoA ligase homology <ACL>

Query Match 64.7%; Score 44; DB 2; Length 541;
 Best Local Similarity 66.7%; Pred. No. 7.2;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LDDLEHOGGHV 12
:|||||: 11
Db 349 MDDLEHRTGV 360

RESULT 3
G83712
hypothetical protein BH0503 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C:Accession: G83712
R:Takami, H.; Nakasone, K.; Takaki, Y.; Mieno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20263314
A:Accession: G83712
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <STM>
A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA04222.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0503
C:Superfamily: Deinococcus radiodurans hypothetical protein DR0763

Query Match 63.2%; Score 43; DB 2; Length 174;
Best Local Similarity 58.3%; Pred. No. 3.1;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LDDLEHOGGHV 12
|:|:|:|:|:|:|:
Db 88 LNDWLHRRGHT 99

RESULT 4
S01433
repressor protein C - phage phi-C31
C:Species: phage phi-C31
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 04-Mar-2000
C:Accession: S01433; S38912
R:Sinclair, R.B.; Bibb, M.J.
Mol. Gen. Genet. 213, 269-277, 1988
A:Title: The repressor gene (C) of the Streptomyces temperate phage phi-C31: nucleotide
A:Reference number: S01433; MUID:89039715
A:Accession: S01433
A:Molecule type: DNA
A:Residues: 1-683 <STM>
A:Cross-references: EMBL:X12865; NID:g15458; PIDN:CAA31345.1; PID:g15459
R:Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Chater, K.F.
Submitted to the EMBL Data Library, November 1993
A:Reference number: S38912
A:Accession: S38912
A:Molecule type: DNA
A:Residues: 1-683 <HAR>
A:Cross-references: EMBL:X76288; NID:g432610; PIDN:CAA53911.1; PID:g432611
C:Genetics:
A:Gene: C
C:Superfamily: phage phi-C31 repressor protein C
C:Keywords: DNA binding; transcription regulation

Query Match 63.2%; Score 43; DB 2; Length 683;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHOGGH 11
|||:|:|:|:|:
Db 474 DDDVEROGGH 483

RESULT 5
E82750
hemagglutinin-like secreted protein XF0889 [imported] - Xylella fastidiosa (strain 9a
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82750
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3282 <STM>
A:Cross-references: GB:AE003928; GB:AE003849; NID:g9105798; PIDN:AAFB3699.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, C.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0889

Query Match 61.8%; Score 42; DB 2; Length 3282;
Best Local Similarity 41.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDDLEHOGGHV 12
|:|:|:|:|:|:|:
Db 472 LENDIDNRGHT 483

RESULT 6
E82589
hemagglutinin-like secreted protein XF2196 [imported] - Xylella fastidiosa (strain 9a
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82589
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82589
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3442 <STM>
A:Cross-references: GB:AE004032; GB:AE003849; NID:g9107324; PIDN:AAFB4995.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, C.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.

A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2196

Query Match 61.8%; Score 42; DB 2; Length 3442;
Best Local Similarity 41.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDDDLHQGHV 12
|:::|:::|:::
Db 472 LENDIDNRGHI 483

RESULT 7
B82519
hemagglutinin-like secreted protein XF2775 [Imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82519

Ranonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82519

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-3455 <SIM>

A:Cross-References: GB:AE004082; GB:AE003849; NID:g9108003; PIDN:AAF5560.1; GSPDB:GN001
A:Experimental source: strain 9a5c

R. Simpson, A.J.G.; Rehnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brites, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carriro, D.M.; Carrer, H
as Neto, E.; Docena, C.; El-Porty, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, F
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sanelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2775

Query Match 61.8%; Score 42; DB 2; Length 3455;
Best Local Similarity 41.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDDDLHQGHV 12
|:::|:::|:::
Db 472 LENDIDNRGHI 483

RESULT 8
O4CKA4
cytochrome P450 52A4 - yeast (Candida maltosa)
N:Alternate names: cytochrome P450-cm2
N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: Candida maltosa
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 03-Mar-2000
C:Accession: S08668; A56578; S65523

R.Schunck, W.H.; Gross, B.; Mueller, H.G.

submitted to the EMBL Data Library, February 1990

A:Reference number: S08667

A:Accession: S08668

A:Molecule type: mRNA
A:Residues: 1-538 <SCH>

A:Cross-References: EMBL:X51932; NID:g2609; PIDN:CAA36198.1; PID:g2610

A:Experimental source: strain EH15D
R.Schunck, W.H.; Vogel, F.; Gross, B.; Kargel, E.; Mauersberger, S.; Kopke, K.; Gengen
Eur. J. Cell Biol. 55, 336-345, 1991

A:Title: Comparison of two cytochromes P-450 from Candida maltosa: primary structures
endoplasmic reticulum.

A:Reference number: A56578; MUID:92037671

A:Accession: A56578

A:Molecule type: mRNA

A:Residues: 1-255, 'D', 257-538 <SC2>

A:Note: sequence inconsistent with nucleotide translation

A:Note: sequence extracted from NCBI backbone (NCBIN:64322, NCBI:64324)

R.Scheller, U.; Zimmer, T.; Kaergel, E.; Schunck, W.H.

Arch. Biochem. Biophys. 328, 245-254, 1996

A:Title: Characterization of the n-alkane and fatty acid hydroxylating cytochrome P45

A:Reference number: S65522; MUID:96213873

A:Accession: S65523

A:Molecule type: protein

A:Residues: 2-15 <SCW>

C:Genetics:

A:Gene: CYP52A4

C:Superfamily: Candida cytochrome P450 52A3; cytochrome P450 homology

C:Keywords: Chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase

F:2-538/Product: cytochrome P450 52A4 #status experimental <MRT>

F:331-507/Domain: cytochrome P450 homology <P45>

F:485/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 60.3%; Score 41; DB 1; Length 538;
Best Local Similarity 72.7%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDLEHQGHV 12
|:::|:::|:::
Db 299 DDDLEHQGHV 309

RESULT 9

T27668

hypothetical protein ZK1053.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T27668

R.Kershaw, J.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20402

A:Accession: T27668

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-536 <WIL>

A:Cross-References: EMBL:Z82084; PIDN:CAB04976.1; GSPDB:GN00019; CESP:ZK1053.2

A:Experimental source: clone ZK1053

C:Genetics:

A:Gene: CESP:ZK1053.2

A:Map position: 1

A:introns: 56/3; 88/3; 203/1; 228/2; 383/1; 467/3

C:Superfamily: Caenorhabditis elegans hypothetical protein ZK1053.2

Query Match 59.6%; Score 40.5; DB 2; Length 536;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 4; Mismatches 0; Indels 3; Gaps 1;

QY 1 LDDDLHQ---GCHV 12
|:::|:::|:::
Db 511 LDDVDHKKRRGGYV 525

RESULT 10

S69790

fibronectin-binding protein II precursor - Streptococcus pyogenes

C:Species: Streptococcus pyogenes

C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 15-Oct-1999

C:Accession: S69790; S72666
 R:Reikemeyer, B.; Talay, S.R.; Chhatwal, G.S.
 M.O.L. Microbiol. 17, 137-145, 1995
 A:Title: Characterization of a novel fibronectin-binding surface protein in group A streptococcus
 A:Reference number: S69790; MUID:96020668
 A:Accession: S69790
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1025 <KRE>
 A:Cross-references: EMBL:X83303; NID:q1070387; PIDN:CAA58282.1; PID:q1552714
 A:Experimental source: strain A75
 A>Note: the authors translated the initiation codon TTG for residue 1 as Leu
 R:Reikemeyer, B.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S72666
 A:Accession: S72666
 A:Molecule type: DNA
 A:Residues: 1-19, 'T', 21-145, 'T', 147-1025 <KRM>
 A:Cross-references: EMBL:X83303; NID:q1070387; PIDN:CAA58282.1; PID:q1552714
 C:Genetics:
 A:Gene: sfbII
 A:Start codon: TTG
 C:Keywords: fibronectin binding; membrane bound
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-1025/Product: fibronectin-binding protein II #status predicted <MAT>
 F:848-868/Region: fibronectin binding
 F:869-907/Region: fibronectin binding
 F:908-946/Region: fibronectin binding
 F:990-994/Region: membrane anchor cleavage motif
 F:999-1016/Domain: transmembrane #status predicted <TM>

Query Match 58.8%; Score 40; DB 2; Length 1025;
 Best Local Similarity 70.0%; Pred. No. 72;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 DDDLEHOGGH 11
 |||:|:|
 DB 809 DDDLEHOGGH 818

RESULT 11
 S27939
 tensin - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 12-Feb-1999
 C:Accession: S27939; S28973
 R:Chen, L.B.
 submitted to the EMBL Data Library, August 1991
 A:Reference number: S27939
 A:Accession: S27939
 A:Molecule type: mRNA
 A:Residues: 1-1733 <CHE>
 A:Cross-references: EMBL:M74165; NID:q212751; PID:q212752
 R:Weygt, C.; Gaertner, A.; Wegner, A.; Korte, H.; Meyer, H.E.
 J. Mol. Biol. 227, 593-595, 1992
 A:Title: Occurrence of an actin-inserting domain in tensin.
 A:Reference number: S28973; MUID:93021103
 A:Accession: S28973
 A:Molecule type: protein
 A:Residues: 862-871, 'X', 873-875, 'A', 877-1212 <WEI>
 C:Superfamily: SH2 homology
 F:1461-1570/Domain: SH2 homology <SH2>

Query Match 58.8%; Score 40; DB 2; Length 1733;
 Best Local Similarity 63.6%; Pred. No. 1.3e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LDDLEHOGGH 11
 |||:|:|
 DB 568 LDDLEHOGGH 578

RESULT 12
 A54970
 tensin, cardiac muscle - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 21-Jul-2000
 C:Accession: A54970; S38330; S21544
 R:Lo, S.H.; An, O.; Bao, S.; Wong, W.K.; Liu, Y.; Janney, P.A.; Hartwig, J.H.; Chen, J.
 J. Biol. Chem. 269, 22310-22319, 1994
 A:Title: Molecular cloning of chick cardiac muscle tensin. Full-length cDNA sequence.
 A:Reference number: A54970; MUID:94350987
 A:Accession: A54970
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-1744 <LOA>
 A:Cross-references: GB:M96625
 R:van de Werken, R.; Gennart, M.; Taveilla, S.; Bet, P.; Molina, F.; Lin, S.; Canceda
 Eur. J. Biochem. 217, 781-790, 1993
 A:Title: Modulation of tensin and vimentin expression in chick embryo developing cartilage
 A:Reference number: S38330; MUID:94039118
 A:Accession: S38330
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1469-1744 <VAN>
 A:Cross-references: EMBL:X66286; NID:q63802; PIDN:CAA46992.1; PID:q63803
 C:Superfamily: SH2 homology
 C:Keywords: cardiac muscle; heart
 F:1472-1581/Domain: SH2 homology <SH2>

Query Match 58.8%; Score 40; DB 2; Length 1744;
 Best Local Similarity 63.6%; Pred. No. 1.3e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LDDLEHOGGH 11
 |||:|:|
 DB 567 LDDLEHOGGH 577

RESULT 13
 A57075
 tensin - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 21-Jul-2000
 C:Accession: A57075
 R:Chuang, J.Z.; Lin, D.C.; Lin, S.
 J. Cell Biol. 128, 1095-1109, 1995
 A:Title: Molecular cloning, expression, and mapping of the high affinity actin-cappin
 A:Reference number: A57075; MUID:95204530
 A:Accession: A57075
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1792 <CHU>
 A:Cross-references: GB:L06662; NID:q212754; PIDN:AA73949.1; PID:q212755
 C:Superfamily: SH2 homology
 F:1520-1629/Domain: SH2 homology <SH2>

Query Match 58.8%; Score 40; DB 2; Length 1792;
 Best Local Similarity 63.6%; Pred. No. 1.3e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LDDLEHOGGH 11
 |||:|:|
 DB 615 LDDLEHOGGH 625

RESULT 14
 F83503
 hypothetical protein PA1130 [Imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: F83503

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lam,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337
 A:Accession: F83503
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <STO>
 A:Cross-references: GB:AE004543; GB:AE004091; NID:g9947047; PIDN:AA04519.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA1130

Query Match 57.4%; Score 39; DB 2; Length 325;
 Best Local Similarity 87.5%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 DDDLEHOG 9
 Db 17 DDDLEHOG 24

RESULT 15

A48529

mitochondrial processing peptidase (EC 3.4.24.64) 55K protein precursor - potato

N:Alternate names: P-35; ubiquinol--cytochrome-c reductase (EC 1.10.2.2) core protein I

C:Species: Solanum tuberosum (potato)

C>Date: 28-Mar-1994 #sequence_revision 14-Jul-1994 #text_change 31-Dec-2000

C:Accession: A48529

R:Emmermann, M.; Braun, H.P.; Arreitz, M.; Schmitz, U.K.

J. Biol. Chem. 268, 18936-18942, 1993

A>Title: Characterization of the bifunctional cytochrome c reductase-processing peptidas

A:Reference number: A48529; MUID:93366812

A:Accession: A48529

A:Molecule type: mRNA

A:Residues: 1-534 <EMM>

A:Experimental source: var. Marfona, tuber

A>Note: sequence extracted from NCBI backbone (NCBIP:136740)

C:Superfamily: mitochondrial processing peptidase alpha chain

C:Keywords: hydrolase; metalloproteinase; mitochondrial matrix; mitochondrion; oxidative

F:1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>

F:33-534/Product: mitochondrial processing peptidase 55K protein #status experimental <M

Query Match 57.4%; Score 39; DB 1; Length 534;
 Best Local Similarity 41.7%; Pred. No. 52;

Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 LDDLEHOGHV 12
 Db 161 LDEEINMGH 172

Search completed: July 6, 2001, 09:17:58
 Job time: 644 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: July 6, 2001, 09:26:36 ; Search time 37.59 Seconds

(without alignments)
10.936 Million cell updates/sec

Title: US-09-437-912-2

Sequence: 1 LDDDLHGGSHV 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	43	63.2	683	1 RPC_BPHC	P08979 bacterioph
3	41	60.3	538	1 CP5D_CANMA	P16141 candida mal
4	40	58.8	1744	1 TENS_CHICK	O04205 gallus gall
5	39	57.4	551	1 CBX4_MOUSE	O55187 mus musculu
6	39	57.4	558	1 CBX4_HUMAN	O00257 homo sapien
7	39	57.4	570	1 DID_ECOLI	P06149 escherichia
8	39	57.4	651	1 BGLR_HUMAN	P08236 homo sapien
9	39	57.4	744	1 RELA_ECOLI	P15585 escherichia
10	38	55.9	573	1 GLB1_MAIZE	P15590 zea mays (m
11	38	55.9	818	1 NAHL_PIG	P48762 sus scrofa
12	37.5	55.1	511	1 SCRB_ZYMO	P22632 zymomonas m
13	37	54.4	426	1 ORDL_ECOLI	P37906 escherichia
14	37	54.4	648	1 BGLR_MOUSE	P12265 mus musculu
15	37	54.4	648	1 BGLR_MOUSE	P06760 rattus norv
16	36.5	53.7	512	1 INVA_ZYMO	P35636 zymomonas m
17	36	52.9	239	1 2138_HUMAN	P52744 homo sapien
18	36	52.9	273	1 TC1A_CABER	P35072 caenorhabdi
19	36	52.9	341	1 ETFA_SCHPO	P78790 schizosacch
20	36	52.9	354	1 ALKB_ARATH	O9a898 arabidopsis
21	36	52.9	430	1 MS12_AGRH	P50201 agrobacteri
22	36	52.9	442	1 KRE2_YEAS	P27809 saccharomyc
23	36	52.9	651	1 BGLR_CANFA	O18835 canis famli
24	36	52.9	651	1 BGLR_CANFA	O95524 felis silve
25	36	52.9	827	1 Y23_METUA	P40282 methanococc
26	36	52.9	844	1 SECA_STACA	P47994 staphylococ
27	36	52.9	906	1 CTN1_MOUSE	P26231 mus musculu
28	35	51.5	182	1 GVH1_HALNI	P24372 halobacteri
29	35	51.5	273	1 PYR2_SYNY3	P73204 synecocyst
30	35	51.5	288	1 YHCS_HAELN	P30311 haemophilus
31	35	51.5	316	1 YOVA_CAEEL	P30634 caenorhabdi
32	35	51.5	389	1 TRP1_MAIZE	P43283 zea mays (m
33	35	51.5	494	1 HMBC_DROME	P09081 drosophila

ALIGNMENTS

RESULT 1	ID	KNG_HUMAN	STANDARD:	PRT:	644 AA.
AC	P01042	P01043			
DT	21-JUL-1986	(Rel. 01, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].				
CN	KNG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RE	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).				
RC	TISUE=Liver;				
RX	MEDLINE=85234582; PubMed=2989293;				
RA	Takagaki Y., Kitamura N., Nakanishi S.;				
RT	"Cloning and sequence analysis of cDNAs for human high molecular weight and low molecular weight prekinnogens. Primary structures of two human prekinnogens."				
RT	J. Biol. Chem. 260:8601-8609(1985).				
RL	[2]				
RN	GENE STRUCTURE.				
RP	MEDLINE=85234583; PubMed=2989294;				
RA	Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,				
RT	Nakanishi S.;				
RT	"Structural organization of the human kininogen gene and a model for its evolution."				
RT	J. Biol. Chem. 260:8610-8617(1985).				
RL	[3]				
RN	SEQUENCE OF 1-401 FROM N.A.				
RP	MEDLINE=85122621; PubMed=6441591;				
RA	Ohnubo I., Kurauchi K., Takasawa T., Shikawa H., Sasaki M.;				
RT	"Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kininogen."				
RT	Biochemistry 23:5691-5697(1984).				
RL	[4]				
RN	SEQUENCE OF 379-644.				
RP	MEDLINE=86030270; PubMed=4054110;				
RA	Lottspeich F., Kellermann J., Henschen A., Foerisch B.,				
RT	Mueller-Esterl W.;				
RT	"The amino acid sequence of the light chain of human high-molecular-mass kininogen."				
RT	Eur. J. Biochem. 152:307-314(1985).				
RL	[5]				
RN	SEQUENCE OF 381-389.				
RP	MEDLINE=90255622; PubMed=4952632;				
RA	Pierce J.V.;				
RT	"Structural features of plasma kinins and kininogens."				
RT	Fed. Proc. 27:52-57(1968).				
RL	[6]				
RN	DISULFIDE BONDS.				
RP	Sueyoshi T., Miyata T., Kato H., Iwanaga S.;				
RA	"Disulfide bonds in bovine HMW kininogens."				
RT					

34	35	51.5	528	1	BGL_ARATH	O9a50 arabidopsis
35	35	51.5	643	1	NOZ_ALCEU	O59105 alcaligenes
36	35	51.5	906	1	CTN1_HUMAN	P35221 homo sapien
37	35	51.5	1612	1	ATCA_YEAST	O12675 saccharomyc
38	34	50.0	163	1	GVH2_HALNI	O9hht6 halobacteri
39	34	50.0	163	1	GVH2_HALNI	P33961 halobacteri
40	34	50.0	316	1	PAXX_ECOLI	P76086 escherichia
41	34	50.0	323	1	ARG2_SCHPO	O10066 schizosacch
42	34	50.0	357	1	DHAS_STRMO	P10539 streptococc
43	34	50.0	362	1	COTB_BACSU	O45535 bacillus su
44	34	50.0	379	1	PSI_SCHPO	O09912 schizosacch
45	34	50.0	390	1	MET1_HUMAN	O00470 homo sapien

RL Seikagaku 56:808-808(1984).

CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)

CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY

CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO

CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-

CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE

CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF

CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE

CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND

CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A

CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR

CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF

CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS

CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,

CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)

CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-

CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD

CC CLOTTING.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: K02566; AAA35497.1; -

DR EMBL: M11437; AAB59550.1; JOINED.

DR EMBL: M11438; AAB59550.1; JOINED.

DR EMBL: M11521; AAB59550.1; JOINED.

DR EMBL: M11522; AAB59550.1; JOINED.

DR EMBL: M11523; AAB59550.1; JOINED.

DR EMBL: M11524; AAB59550.1; JOINED.

DR EMBL: M11525; AAB59550.1; JOINED.

DR EMBL: M11526; AAB59550.1; JOINED.

DR EMBL: M11527; AAB59550.1; JOINED.

DR EMBL: M11528; AAB59550.1; JOINED.

DR EMBL: M11437; AAB59551.1; JOINED.

DR EMBL: M11438; AAB59551.1; JOINED.

DR EMBL: M11521; AAB59551.1; JOINED.

DR EMBL: M11522; AAB59551.1; JOINED.

DR EMBL: M11523; AAB59551.1; JOINED.

DR EMBL: M11524; AAB59551.1; JOINED.

DR EMBL: M11525; AAB59551.1; JOINED.

DR EMBL: M11526; AAB59551.1; JOINED.

DR EMBL: M11527; AAB59551.1; JOINED.

DR EMBL: M11528; AAB59551.1; JOINED.

DR PIR: A01279; KGH01.

DR PIR: A25276; A25276.

DR PIR: A01280; KGH01.

DR PIR: B25276; B25276.

DR PIR: S02482; S02482.

DR SWISS-2DPAGE; P01043; HUMAN.

DR MIM; 228960; -

DR InterPro: IPR00010; -

DR InterPro: IPR002395; -

DR Pfam: PF00031; Cystatin; 3.

DR PRINTS; PR00334; KININOGEN.

DR PROSITE; PS00287; CYSTATIN; 2.

KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;

KW Bradykinin; Blood coagulation; Inflammatory response; Signal;

KW Alternative splicing

FT SIGNAL 1 18

FT CHAIN 19 644 KININOGEN.

FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 389 BRADYKININ.

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD_RES	19	19	
FT	DISULFID	28	614	
FT	DISULFID	83	94	
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	205	205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	294	294	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPPLIC	402	427	VSPPHSMAPADDERDGGKRGHTR -> SHLRSCXEYKGR
FT	VARSPPLIC	428	644	PKKAEAPASEREYS (IN ISOFORM LMW).
FT	CONFLICT	593	593	T -> I (IN REF. 1).
SO	SEQUENCE	644 AA; 71945 MW;	3132BACBAF8FB87E CRC64;	

Query Match 100.0%; Score 68; DB 1; Length 644;

Best Local Similarity 100.0%; Pred. No. 0.00032;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDDDLEHOGCHV 12

DB 479 LDDDDLEHOGCHV 490

RESULT 2

RPC_BPPHC STANDARD; PRT; 683 AA.

AC P08379; 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE REPRESSOR PROTEIN C.

GN C.

OS Bacteriophage phi-C31.

OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;

OC Lambda phage group.

OX NCBI_Taxid=10719;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C+-NORWICH;

RX MEDLINE=89039715; PubMed=3185504;

RA Sinclair R.B., Bibb M.J.;

RT "The repressor gene (c) of the Streptomyces temperate phage phi c31:

RT Nucleotide sequence, analysis and functional cloning.";

RL Mol. Gen. Genet. 213:269-277(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94374705; PubMed=8088546;

RA Hartley N.M., Murphy G.O., Bruton C.J., Chater K.F.;

RT "Sequence of the essential early region of phi C31, a temperate phage
RT of Streptomyces spp. with unusual features in its lytic
RT development.";
RL Gene 147:29-40(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X12865; CAA31345.1; -
DR EMBL: X76288; CAA53911.1; -
DR PIR: S01433; S01433.
KM Transcription regulation; Repressor; DNA-binding.
SO SEQUENCE 683 AA; 74077 MW; B02379D204F37D1B CRC64;

Query Match 63.2%; Score 43; DB 1; Length 683;
Best Local Similarity 70.0%; Pred. No. 6.7;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGGH 11
|||:|:|:|
DB 474 DDDVEROGGH 483

RESULT 3
CP5D_CANMA STANDARD; PRT; 538 AA.
ID CP5D_CANMA
AC P16141.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 52A4 (EC 1.14.14.1) (CYPL1TA4) (ALKANE-INDUCIBLE P450-
DE ALK-A) (P450-CM2).
CN CYP52A4.
OS Candida maltosa (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5479;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EH15D;
RX MEDLINE=89286595; PubMed=2735924;
RA Schunck W.-H., Kaerger E., Gross B., Wiedmann B., Mauersberger S.,
RA Koepke K., Klessing U., Strauss M., Gaestel M., Mueller H.-G.;
RT "Molecular cloning and characterization of the primary structure of
RT the alkane hydroxylating cytochrome P-450 from the yeast Candida
RT maltosa."
RT Blochem. Biophys. Res. Commun. 161:843-850(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91229697; PubMed=2039569;
RA Ohkuma M., Tanimoto T., Yano K., Takagi M.;
RT "CYP52 (cytochrome P450alk) multigene family in Candida maltosa:
RT molecular cloning and nucleotide sequence of the two tandemly
RT arranged genes."
RT DNA Cell Biol. 10:271-282(1991).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=96311366; PubMed=8713123;
RA Zimmer T., Ohkuma M., Ohta A., Takagi M., Schunck W.H.;
RT "The CYP52 multigene family of Candida maltosa encodes functionally
RT diverse n-alkane-inducible cytochromes P450."
RT Biochem. Biophys. Res. Commun. 224:784-789(1996).
CC -1- FUNCTION: TOGETHER WITH AN NADPH CYTOCHROME P450 THE ENZYME SYSTEM
CC CATALYZES THE TERMINAL HYDROXYLATION AS THE FIRST STEP IN THE
CC ASSIMILATION OF ALKANES AND FATY ACIDS.
CC -1- CATALYTIC ACTIVITY: HYDROXYLATION OF N-ALKANES AT THE TERMINAL
CC POSITION.

CC -1- INDUCTION: BY ALKANES.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51932; CAA36198.1; -
DR PIR: S08668; O4CKA4.
DR InterPro: IPR001128; -
DR InterPro: IPR002402; -
DR InterPro: IPR002974; -
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PRINTS: PR01239; EP450IICYP52.
DR PRINTS: PR01239; EP450IICYP52.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Electron transport; Oxidoreductase; Monooxygenase; Heme;
KW Transmembrane.
FT BINDING 27 46 HEME (BY SIMILARITY).
FT CONFLICT 2 2 S -> P (IN REF. 2).
FT CONFLICT 39 39 P -> N (IN REF. 2).
FT CONFLICT 235 235 V -> T (IN REF. 2).
FT CONFLICT 299 300 DD -> EA (IN REF. 2).
FT CONFLICT 442 442 S -> L (IN REF. 2).
FT CONFLICT 449 449 S -> N (IN REF. 2).
FT CONFLICT 514 514 N -> D (IN REF. 2).
FT CONFLICT 527 527 V -> L (IN REF. 2).
SQ SEQUENCE 538 AA; 61844 MW; 3F540F89C793F03B CRC64;

Query Match 60.3%; Score 41; DB 1; Length 538;
Best Local Similarity 72.7%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGGHV 12
|||:|:|:|
DB 299 DDDLEKOGGYV 309

RESULT 4
TENS_CHICK STANDARD; PRT; 1744 AA.
ID TENS_CHICK
AC 004205; Q91007; Q92011;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TENSIN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=94350987; PubMed=8071358;
RA Lo S.H., An Q., Bao S., Wong W.K., Liu Y., Janney P.A., Hartwig J.H.,
RA Chen L.B.;
RT "Molecular cloning of chick cardiac muscle tensin. Full-length cDNA
RT sequence, expression, and characterization."
RT J. Biol. Chem. 269:22310-22319(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Heart;
RX MEDLINE=95204530; PubMed=7896874;
RA Chuang J.Z., Lin D.C., Lin S.;
RT "Molecular cloning, expression, and mapping of the high affinity

RT actin-capping domain of chicken cardiac tensin.";
 RL J. Cell Biol. 128:1095-1109(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chen L.B.;
 RL Submitted (XXX-1991) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1469-1744 FROM N.A.
 RC TISSUE-Embryonic chondrocytes, and Embryonic heart;
 RX MEDLINE-94039118; PubMed-8223621;
 RA van de Werken R., Gennari M., Tavella S., Bet P., Molina F.,
 Lin S., Canceda R., Castagnola P.;
 RT "Modulation of tensin and vimentin expression in chick embryo
 dev. J. Biochem. 217:781-790(1993).
 RN [5]
 RP SH2 DOMAIN
 RX MEDLINE-91220073; PubMed-1708917;
 RA Davis S., Lu M.L., Lo S.H., Lin S., Butler J.A., Druker B.J.,
 Roberts T.M., An Q., Chen L.B.;
 RT "Presence of an sh2 domain in the actin-binding protein tensin.";
 RL Science 253:712-715(1991).
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL MIGRATION, CARTILAGE DEVELOPMENT
 AND IN LINKING SIGNAL TRANSDUCTION PATHWAYS TO THE CYTOSKELETON.
 CC BINDS TO ACTINS AND PHOSPHORYLATED PROTEINS IN SRC-TRANSFORMED
 CC CELLS. MAY BIND ACTIN WITH CAPPING AND BUNDLING PROPERTIES.
 CC -1- SUBCELLULAR LOCATION: LOCALIZES TO ADHERENS JUNCTIONS.
 CC -1- TISSUE SPECIFICITY: HEART, GIZZARD, LUNG AND SKELETAL MUSCLE.
 CC -1- PTM: TYROSINE-PHOSPHORYLATED.
 CC -1- SIMILARITY: CONTAINS 1 TENSIN DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M96625; AAA59053.1; -
 DR EMBL: L06662; AAA73949.1; ALT_INIT.
 DR EMBL: Z18529; CAA79215.1; ALT_INIT.
 DR EMBL: M74165; AAA49087.1; -
 DR EMBL: X66286; CAA46992.1; -
 DR HSSP: P16277; 1BLJ
 DR InterPro: IPR000980; -
 DR Pfam: PF00017; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 KW Actin-binding; Cytoskeleton; SH2 domain; Phosphorylation.
 FT DOMAIN 66 342
 FT 1472 1581 SH2.
 FT CONFLICT 49 49 C-> R (IN REF. 2).
 FT CONFLICT 61 61 M-> T (IN REF. 2).
 FT CONFLICT 88 88 Q-> T (IN REF. 3).
 FT CONFLICT 404 404 A-> T (IN REF. 2).
 FT CONFLICT 452 452 T-> A (IN REF. 2).
 FT CONFLICT 508 509 DV-> EL (IN REF. 2).
 FT CONFLICT 522 522 A-> P (IN REF. 2).
 FT CONFLICT 664 664 R-> A (IN REF. 2).
 FT CONFLICT 666 666 A-> T (IN REF. 2).
 FT CONFLICT 875 875 P-> A (IN REF. 2).
 FT CONFLICT 909 909 P-> T (IN REF. 2).
 FT CONFLICT 1102 1113 MISSING (IN REF. 3).
 FT CONFLICT 1240 1240 G-> A (IN REF. 2).
 FT CONFLICT 1480 1480 E-> D (IN REF. 2).
 FT CONFLICT 1711 1711 D-> E (IN REF. 2).
 SQ SEQUENCE 1744 AA; 187214 MW; 5C3C8B6211935524 CRC64;

OY I LDDLEHOGGH 11
 DB 567 LDDLEPNOGDH 577

RESULT 5
 CBX4_MOUSE STANDARD; PRT: 551 AA.
 ID CBX4_MOUSE
 AC 055187;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CHROMOX PROTEIN HOMOLOG 4 (POLYCOMB 2 HOMOLOG) (PC2) (MPC2).
 GN CBX4 OR MPC2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98035734; PubMed-9367786;
 RA Alkema M.J., Jacobs J., Voncken J.W., Jenkins N.A., Copeland N.G.,
 Satijn D.P.E., Otte A.P., Berns A., van Lohuizen M.;
 RT "Mpc2, a new murine homolog of the Drosophila polycomb protein is a
 RT member of the mouse polycomb transcriptional repressor complex.";
 RL J. Mol. Biol. 273:993-1003(1997).
 CC -1- FUNCTION: INVOLVED IN MAINTAINING THE TRANSCRIPTIONALLY REPRESSIVE
 CC STATE OF GENES. MODIFIES CHROMATIN, RENDERING IT HERITABLY CHANGED
 CC IN ITS EXPRESSIBILITY.
 CC -1- SUBUNIT: COMPONENT OF THE CHROMATIN-ASSOCIATED POLYCOMB COMPLEX
 CC (PCG).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: CONTAINS 1 'CHROMO' DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U63387; AAB96874.1; -
 DR HSSP: P23197; LAP0
 DR MGD: MGI:1195985; Cdx4.
 DR InterPro: IPR000953; -
 DR Pfam: PF00385; Chromo; 1.
 DR PRINTS: PR00504; CHROMODOMAIN.
 DR PROSITE: PS00598; CHROMO_1; 1.
 DR PROSITE: PS50013; CHROMO_2; 1.
 KW Chromatin regulator; Nuclear protein; Transcription regulation;
 KW Repressor.
 FT DOMAIN 11 69
 FT 383 395 CHROMO.
 FT CONFLICT 383 395 POLY-HIS.
 SQ SEQUENCE 551 AA; 60581 MW; 30CEB09A82C58400 CRC64;

Query Match 57.4%; Score 39; DB 1; Length 551;
 Best Local Similarity 75.0%; Pred. No. 26;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 DLEHOGGH 11
 DB 170 DLEHOGGH 177

Query Match 58.8%; Score 40; DB 1; Length 1744;
 Best Local Similarity 63.6%; Pred. No. 62;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 6
 CBX4_HUMAN STANDARD; PRT: 558 AA.
 ID CBX4_HUMAN
 AC 000257;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CHROMOX PROTEIN HOMOLOG 4 (POLYCOMB 2 HOMOLOG) (PC2) (HPC2).
GN CBX4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=97459707; PubMed=9315667;
RA Saitjn D.P.E., Olson D.J., van der Vlag J., Hamer K.M., Lambrechts C.,
RA Masselink H., Gunster M.J., Sewalt R.G.A.B., van Driel R., Otte A.P.;
RT "Interference with the expression of a novel human polycomb protein,
RT hpc2, results in cellular transformation and apoptosis.";
RL Mol. Cell. Biol. 17:6076-6086(1997).
RN [2]
RP SEQUENCE OF 455-558 FROM N.A.
RX MEDLINE=97342649; PubMed=9199346;
RA Saitjn D.P.E., Gunster M.J., van der Vlag J., Hamer K.M., Schul W.,
RA Alkema M.J., Saurin A.J., Freemont P.S., van Driel R., Otte A.P.;
RT "RING1 is associated with the polycomb group protein complex and acts
RT as a transcriptional repressor.";
RL Mol. Cell. Biol. 17:4105-4113(1997).
CC -1- FUNCTION: INVOLVED IN MAINTAINING THE TRANSCRIPTIONALLY REPRESSIVE
CC STATE OF GENES. MODIFIES CHROMATIN, RENDERING IT HERITABLY CHANGED
CC IN ITS EXPRESSIBILITY.
CC -1- SUBUNIT: COMPONENT OF THE CHROMATIN-ASSOCIATED POLYCOMB COMPLEX
CC (PCG).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: UBICUITOUS.
CC -1- SIMILARITY: CONTAINS 1 'CHROMO' DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF013956; AAB80718.1; -
DR EMBL: U94344; AAB62734.1; -
DR MIM: 603079; -
DR HSSP: P23197; IAP0.
DR InterPro: IPR000953; -
DR Pfam: PF00385; Chromo; 1.
DR PRINTS: PR00504; CHROMODOMAIN.
DR PROSITE: PS00598; CHROMO_1; 1.
DR PROSITE: PS50013; CHROMO_2; 1.
KW Chromatin regulator; Nuclear protein; Transcription regulation;
KW Repressor.
FT DOMAIN 16 69 CHROMO.
FT DOMAIN 383 398 POLY-HIS.
FT DOMAIN 499 508 POLY-ALA.
SQ SEQUENCE 558 AA: 61228 MW: 7158526991D33463 CRC64:

Query Match 57.4%; Score 39; DB 1; Length 558;
Best Local Similarity 75.0%; Pred. NO. 26;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DLEHOGGH 11
11:111111
DB 168 DLYOGGH 175

RESULT 7
DID_ECOLI
ID DID_ECOLI STANDARD: PRT: 570 AA.
AC P06149;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE D-LACTATE DEHYDROGENASE (Ec 1.1.1.28).
GN DLD.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85027231; PubMed=6386470;
RA Campbell H.D., Rogers B.L., Young I.G.;
RT "Nucleotide sequence of the respiratory D-lactate dehydrogenase gene
RT of Escherichia coli.";
RL Eur. J. Biochem. 144:367-373(1984).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-17.
RX MEDLINE=85130784; PubMed=3882663;
RA Rule G.S., Pratt E.A., Chin C.C.O., Mold F., Ho C.;
RT "Overproduction and nucleotide sequence of the respiratory D-lactate
RT dehydrogenase of Escherichia coli.";
RL J. Bacteriol. 161:1059-1068(1985).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / BHB2600;
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA Church G.M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=20402548; PubMed=10944213;
RA Dym O., Pratt E.A., Ho C., Eisenberg D.;
RT "The crystal structure of D-lactate dehydrogenase, a peripheral
RT membrane respiratory enzyme.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:9413-9418(2000).
CC -1- FUNCTION: FIRST COMPONENT OF THE MEMBRANE-BOUND D-LACTATE OXIDASE,
CC WHICH IS BELIEVED TO PLAY AN IMPORTANT ROLE IN THE ENERGIZATION OF
CC THE ACTIVE TRANSPORT OF A VARIETY OF SUGARS AND AMINO ACIDS.
CC -1- CATALYTIC ACTIVITY: D-LACTATE + NAD(+) = PYRUVATE + NADH.
CC -1- COFACTOR: FAD.
CC -1- ENZYME REGULATION: REQUIRES PHOSPHOLIPID FOR MAXIMAL ACTIVITY.
CC -1- SUBCELLULAR LOCATION: MEMBRANE BOUND LOCATED AT THE INNER FACE OF
CC THE CYTOPLASMIC MEMBRANE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M10038; AAA23688.1; -
DR EMBL: X01067; CAA25531.1; -
DR EMBL: U00007; AAA60530.1; ALT_INIT.
DR EMBL: AE000302; AAC75194.1; -
DR PIR: A21893; DEECDL.
DR PDB: 1FOX; 23-AUG-00.
DR EC2DDBASE: H062.0; 6TH EDITION.
DR Ecogene: EG10231; did.
DR InterPro: IPR001575; -
DR Pfam: PF01565; FAD_binding_4; 1.
KW Oxidoreductase; NAD; Flavoprotein; FAD; Membrane; 3D-structure.
FT INIT_MET 0 0

SQ SEQUENCE 570 AA: 64481 MW: 930B30CF4405H1A7 CRC64;
 Query Match 57.4%; Score 39; DB 1; Length 570;
 Best Local Similarity 60.0%; Pred. No. 27;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 DDLEHOGGH 11
 111:111
 Db 212 DDVHRDGRH 221
 RESULT 8
 ID BGLR_HUMAN STANDARD; PRT; 651 AA.
 AC P08236;
 DT 01-AUG-1988 (rel. 08, Created)
 DT 01-AUG-1988 (rel. 08, Last sequence update)
 DT 01-OCT-2000 (rel. 40, Last annotation update)
 DE BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31) (BETA-G1).
 GN GUSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87118233; PubMed=3468507;
 RA Oshima A., Klye J.W., Miller R.D., Hoffman J.W., Powell P.P.,
 RT Grubb J.H., Sly W.S., Tropak M., Gulse K.S., Gravel R.A.;
 RT "Cloning, sequencing, and expression of cDNA for human beta-
 RT glucuronidase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:685-689(1987).
 RN [2]
 RP SEQUENCE OF 1-70 FROM N.A.
 RX MEDLINE=92009900; PubMed=1916806;
 RA Shipley J.M., Miller R.D., Wu B.M., Grubb J.H., Christensen S.G.,
 RA Kyle J.W., Sly W.S.;
 RT "Analysis of the 5' flanking region of the human beta-glucuronidase
 RT gene.";
 RL Genomics 10:1009-1018(1991).
 RN [3]
 RP SEQUENCE OF 23-32 AND 160-175.
 RC TISSUE=Placenta;
 RX MEDLINE=92162201; PubMed=1311180;
 RA Tanaka J., Gasa S., Sakurada K., Miyazaki T., Kasai M., Makita A.;
 RT "Characterization of the subunits and sugar moiety of human placental
 RT and leukemic beta-glucuronidase.";
 RL Biol. Chem. Hoppe-Seyler 373:57-62(1992).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=96185449; PubMed=8599764;
 RA Jain S., Drendel W.B., Chen Z.W., Mathews F.S., Sly W.S., Grubb J.H.;
 RT "Structure of human beta-glucuronidase reveals candidate lysosomal
 RT targeting and active-site motifs.";
 RL Nat. Struct. Biol. 3:375-381(1996).
 RN [5]
 RP VARIANT MPS-VII TRP-216.
 RX MEDLINE=94154730; PubMed=8111412;
 RA Vervoort R., Lissens W., Liebaers I.;
 RT "Molecular analysis of a patient with hydrops fetalis caused by beta-
 RT glucuronidase deficiency, and evidence for additional pseudogenes.";
 RL Hum. Mutat. 2:443-445(1993).
 RN [6]
 RP VARIANTS MPS-VII VAL-354 AND TRP-611.
 RX MEDLINE=94154731; PubMed=8111413;
 RA Wu B.M., Sly W.S.;
 RT "Mutational studies in a patient with the hydrops fetalis form of
 RT mucopolysaccharidosis type VII.";
 RL Hum. Mutat. 2:446-457(1993).
 RN [7]
 RP VARIANTS MPS-VII CYS-382 AND VAL-619.
 RX MEDLINE=91090114; PubMed=1702266;

RA Tomatsu S., Fukuda S., Sukegawa K., Ikeda Y., Yamada S., Yamada Y.,
 RA Sasaki T., Okamoto H., Kuwahara T., Yamaguchi S., Kiman T.,
 RA Shintaku H., Ishiki G., Orii T.;
 RT "Mucopolysaccharidosis type VII: characterization of mutations and
 RT molecular heterogeneity.";
 RL Am. J. Hum. Genet. 48:89-96(1991).
 RN [8]
 RP VARIANT MPS-VII CYS-627.
 RX MEDLINE=93190983; PubMed=7680524;
 RA Shipley J.M., Klinkenberg M., Wu B.M., Bachinsky D.R., Grubb J.H.,
 RA Sly W.S.;
 RT "Mutational analysis of a patient with mucopolysaccharidosis type
 RT VII, and identification of pseudogenes.";
 RL Am. J. Hum. Genet. 52:517-526(1993).
 CC -1- FUNCTION: BETA-GLUCURONIDASE PLAYS AN IMPORTANT ROLE IN THE
 CC DEGRADATION OF DERMATAN AND KERATAN SULFATES.
 CC -1- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O = AN
 CC ALCOHOL + D-GLUCURONATE.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.
 CC -1- PTM: GLYCOSYLATED WITH 3 TO 4 N-LINKED OLIGOSACCHARIDE CHAINS.
 CC -1- DISEASE: DEFECTS IN GUSB ARE THE CAUSE OF MUCOPOLYSACCHARIDOSIS
 CC TYPE VII (MPS-VII) (ALSO KNOWN AS SLX SYNDROME).
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M15182; AAA52561.1; -;
 DR EMBL: M65002; AAA52622.1; -;
 DR PIR: A26581; A26581.
 DR PDB: 1BHG; 17-SEP-97.
 DR MIM: 253220; -;
 DR InterPro: IPR001649; -;
 DR Pfam: PF00703; GLYCO_HYDRO_2; 1.
 DR ProSITE: PS00132; GLYHYDRO2.
 DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
 DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
 DR Hydrolase; Glycosidase; Lysosome; Glycoprotein; Signal;
 KW Mucopolysaccharidosis; Disease mutation; 3D-structure.
 FT SIGNAL 1 22
 FT CHAIN 23 651 BETA-GLUCURONIDASE.
 FT ACT_SITE 451 451 PROTON DONOR.
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 216 216 R->W (IN MPS-VII).
 FT VARIANT 354 354 /FTID=VAR_003196.
 FT VARIANT 382 382 A->V (IN MPS-VII).
 FT VARIANT 382 382 /FTID=VAR_003197.
 FT VARIANT 611 611 R->C (IN MPS-VII).
 FT VARIANT 611 611 /FTID=VAR_003198.
 FT VARIANT 619 619 R->W (IN MPS-VII).
 FT VARIANT 619 619 /FTID=VAR_003199.
 FT VARIANT 627 627 A->V (IN MPS-VII).
 FT VARIANT 627 627 /FTID=VAR_003200.
 FT VARIANT 627 627 W->C (IN MPS-VII).
 FT VARIANT 627 627 /FTID=VAR_003201.
 SQ SEQUENCE 651 AA: 74715 MW: 68A7A1935C9ABBD CRC64;
 Query Match 57.4%; Score 39; DB 1; Length 651;
 Best Local Similarity 60.0%; Pred. No. 31;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 3 DDLEHOGGHV 12
 111111111:

Db 138 DTLEHGGYL 147

```

RESULT 9
RELA_ECOLI STANDARD: PRT: 744 AA.
AC P11585:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GTP PYROPHOSPHOKINASE (EC 2.7.6.5) (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE)
DE (PPGPP SYNTHETASE 1) ((P)PPGPP SYNTHETASE).
GN RELA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=85008481; PubMed-2844820;
RA Metzger S., Dror I.B., Alzenman E., Schreiber G., Toone M.,
RA Friesen J.D., Cashel M., Glaser G.;
RT "The nucleotide sequence and characterization of the relA gene of
RT Escherichia coli."
RL J. Biol. Chem. 263:15699-15704(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: IN EUBACTERIA PPGPP (GUANOSINE 3'-DIPHOSPHATE 5'-
CC DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT
CC COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO
CC CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYZES THE
CC FORMATION OF PPGPP WHICH IS THEN HYDROLYSED TO FORM PPGPP.
CC -1- CATALYTIC ACTIVITY: ATP + GTP = AMP + GUANOSINE 3'-DIPHOSPHATE
CC 5'-TRIPHOSPHATE.
CC -1- PATHWAY: FIRST STEP IN THE METABOLISM OF PPGPP.
CC -1- SIMILARITY: BELONGS TO THE RELA / SPOT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04039; AAA03237.1; -
DR EMBL: U29580; AAA69294.1; -
DR EMBL: AE000362; AAC75826.1; -
DR PIR: A31996; KIEG
DR EcGene: BG10835; relA.
DR InterPro: IPR002912; -
DR Pfam: PF01842; ACT; 1.
KW Transferase; Kinase.
FT CONFLICT 307 307 F -> K (IN REF. 1).
FT SEQUENCE 744 AA; 83875 MW; FA269709F151EF25 CRC64;

```

Query Match 57.4%; Score 39; DB 1; Length 744;
 Best Local Similarity 77.8%; Pred. No. 36;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LDDLEHOG 9
 DB 512 LDDLEHOG 520

```

RESULT 10
GLBI_MAIZE STANDARD: PRT: 573 AA.
AC P15590:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLOBULIN-1 S ALLELE PRECURSOR (GLBI-S) (7S-LIKE).
GN GLBI.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, INBRED LINE VA26;
RA Belanger F.C., Kriz A.L.;
RT "Molecular characterization of the major maize embryo globulin encoded
RT by the GLBI gene."
RL Plant Physiol. 91:636-643(1989).
RN [2]
RP SEQUENCE OF 87-100.
RX MEDLINE=89374022; PubMed-2775172;
RA Kriz A.L.;
RT "Characterization of embryo globulins encoded by the maize Glb
RT genes."
RL Biochem. Genet. 27:239-251(1989).
CC -1- PM: THREE PROTEIN-PROCESSING STEPS OCCUR IN THE FORMATION OF THE
CC MATURE PROTEIN FROM THE PRIMARY TRANSLATION PRODUCT.
CC -1- POLYMORPHISM: THE THREE MOST COMMONLY OCCURRING GLBI ALLELES HAVE
CC THE DESIGNATION L, I, AND S FOR LARGE, INTERMEDIATE, AND SMALL
CC PROTEINS, RESPECTIVELY.
CC -1- SIMILARITY: BELONGS TO THE 7S SEED STORAGE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M24845; AAA33467.1; -
DR HSSP: P50477; ICAM.
DR MaizeDB: 30181; -
DR InterPro: IPR001113; -
DR Pfam: PF00546; Seedstore_7s; 1.
KW Seed storage protein; Signal.
FT SIGNAL 1 18 OR 21 (POTENTIAL).
FT PROPEP 19 86
FT CHAIN 87 573
FT CARBOHYD 349 349 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SEQUENCE 573 AA; 65029 MW; 525ED1D0A062976 CRC64;

```

Query Match 55.9%; Score 38; DB 1; Length 573;
 Best Local Similarity 60.0%; Pred. No. 40;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 DDDLEHOGH 11
 DB 26 DDDLEHOGH 35

RESULT 11
 NAHL_PIG STANDARD: PRT: 818 AA.
 AC P48762;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)

DF 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+), EXCHANGER 1) (NHE-1).
 GN SIC9AI OR NHE1.
 OS Sus scrofa (Pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OK NCBI_TaxID=9823;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92087905; PubMed-1661081;
 RA Reilly R.F., Hildebrandt F., Blaymesderfer D., Sardet C.,
 RA Pouyssegur J., Aronson P.S., Slayman C.W., Igarashi P.,
 RT "cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in
 RT LLC-PK1 renal epithelial cells."
 RL Am. J. Physiol. 261:F1088-F1094(1991).
 CC -1- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
 CC TRANSDUCTION.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- PTM: PHOSPHORYLATED (POSSIBLE).
 CC -1- SIMILARITY: BELONGS TO THE NA(+)/H(+), EXCHANGER FAMILY.
 CC -1- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
 CC DOMAINS IN THE NA(+)/H(+), EXCHANGERS VARY AMONG AUTHORS.
 CC -1- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
 CC TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M89631; AAA31092.1; -
 DR EMBL: S71135; AAB20633.1; -
 DR InterPro: IPR001970; -
 DR Pfam: PF00999; Na.H.Exchanger; 1.
 DR PRINTS: PRO1084; NAHEXCHNGR.
 DR Transmembrane: Glycoprotein; Sodium transport; Transport; Symport;
 KW Transmembrane family; Phosphorylation.
 FT DOMAIN 1 11
 FT DOMAIN 12 31
 FT DOMAIN 102 123
 FT DOMAIN 124 126
 FT TRANSMEM 127 146
 FT DOMAIN 147 158
 FT TRANSMEM 159 179
 FT DOMAIN 180 184
 FT TRANSMEM 185 206
 FT DOMAIN 207 226
 FT TRANSMEM 227 247
 FT DOMAIN 248 256
 FT TRANSMEM 257 278
 FT DOMAIN 279 297
 FT TRANSMEM 298 318
 FT DOMAIN 319 332
 FT TRANSMEM 333 353
 FT DOMAIN 354 406
 FT TRANSMEM 385 406
 FT DOMAIN 407 412
 FT TRANSMEM 413 434
 FT DOMAIN 435 448
 FT TRANSMEM 449 469
 FT DOMAIN 470 478
 FT TRANSMEM 479 499
 FT DOMAIN 500 818
 FT CARBOHYD 370 370
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 683 683 H -> Y (IN AAB20633).
 SQ SEQUENCE 818 AA; 90987 MW; 9329F7D9A51D3DC9 CRC64;
 Query Match
 Best Local Similarity 55.9%; Score 38; DB 1; Length 818;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 2 DDLEHOGGHV 12
 DB 756 DDEHDHGGGLV 766
 RESULT 12
 SCRB_ZYMO STANDARD; PRT; 511 AA.
 ID SCRB_ZYMO
 AC P22632;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 DE SUCROSE-6-PHOSPHATE HYDROLASE (EC 3.2.1.26) (SUCRASE) (INVERTASE).
 GN SACA.
 OS Zymomonas mobilis.
 CC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
 CC Zymomonas.
 OK NCBI_TaxID=542;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 10988 / ZM1;
 RX MEDLINE-91072217; PubMed-2254250;
 RA Gunasekaran P., Karunakaran T., Cami B., Mukundan A.G., Preziosi L.,
 RA Bartlett J.;
 RT "Cloning and sequencing of the sacA gene: characterization of a
 RT sucrose from Zymomonas mobilis.";
 RL J. Bacteriol. 172:6727-6735(1990).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-
 CC FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M62718; AAB27701.1; -
 DR PIR: A37803; A37803.
 DR InterPro: IPR01362; -
 DR Pfam: PF00251; Glyco_hydro_32; 1.
 DR PROSITE: PS00609; GLYCOSYL_HYDROL_F32; 1.
 KW Hydrolase; Glycosidase.
 FT ACT_SITE 43 43 BY SIMILARITY
 SQ SEQUENCE 511 AA; 58397 MW; 8020D476C835FA0 CRC64;
 Query Match
 Best Local Similarity 55.1%; Score 37.5; DB 1; Length 511;
 Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 OY 3 DDLEHOG-GHV 12
 DB 185 DDKKHGGIGHV 195
 RESULT 13
 ORDL_ECOLI STANDARD; PRT; 426 AA.
 ID ORDL_ECOLI
 AC P37906;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PROBABLE OXIDOREDUCTASE ORDL (EC 1.-.-.-).
 GN ORDL.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 OX NCBI_TaxID=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Jovanovic G.;
 RL Submitted (Oct-1995) to the EMBL/Genbank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Aliba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiiuchi T.,
 Ikemoto K., Inada T., Isono K., Itoh T., Kanai K., Kasai H.,
 Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,
 Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
 Masumoto H., Nishio Y., Oshima T., Saito N., Sampa G., Seki Y.,
 Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;
 RA Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
 [4]
 RP SEQUENCE OF 1-251 FROM N.A.
 RX MEDLINE=91216440; PubMed=1840553;
 RA Helm R., Strehler E.B.;
 RT "Cloning an Escherichia coli gene encoding a protein remarkably
 RT similar to mammalian aldehyde dehydrogenases.";
 RL Gene 99:15-23(1991).
 [5]
 RP IDENTIFICATION
 RX MEDLINE=95075659; PubMed=7984428;
 RA Borodovsky M., Rudd K.E., Koonin E.V.;
 RT "Intrinsic and extrinsic approaches for detecting genes in a
 RT bacterial genome.";
 RL Nucleic Acids Res. 22:4756-4767(1994).
 CC -1- SIMILARITY: STRONG, TO H.INFLUENZA ORDL.
 CC -1- SIMILARITY: SOME, TO NAD-BINDING SITES OF DEHYDROGENASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U38543; AAC5300.1; -
 DR EMBL: AE000226; AAC74383.1; -
 DR EMBL: D90768; BAA14870.1; -
 DR EMBL: M38433; AAA23429.1; -
 DR Ecogene: EG11822; ordl.
 KW Oxidoreductase; NAD.
 SO SEQUENCE 426 AA; 47169 MW; 44158FF418E9C254 CRC64;

Query Match 54.4%; Score 37; DB 1; Length 426;
 Best Local Similarity 62.5%; Pred. No. 42;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 LEHOGSHV 12
 1:1111:
 Db 174 LDHSGSHI 181

RESULT 14
 ID BGLR_MOUSE STANDARD: PRT; 648 AA.
 AC P12265; Q61601; Q64473; Q64474;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31).
 GN GUSB OR GUS OR GUS-S.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88085188; PubMed=2891607;
 RA Gallagher P.M., D'Amore M.A., Lund S.D., Elliott R.W., Pazik J.,
 RA Holman C., Korfhagen T.R., Ganschow R.E.;
 RT "DNA sequence variation within the beta-glucuronidase gene complex
 RT among inbred strains of mice.";
 RL Genomics 1:145-152(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88284700; PubMed=3397060;
 RA Gallagher P.M., D'Amore M.A., Lund S.D., Ganschow R.E.;
 RT "The complete nucleotide sequence of murine beta-glucuronidase mRNA
 RT and its deduced polypeptide.";
 RL Genomics 2:215-219(1988).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89062453; PubMed=3196706;
 RA D'Amore M.A., Gallagher P.M., Korfhagen T.R., Ganschow R.E.;
 RT "Complete sequence and organization of the murine beta-glucuronidase
 RT gene.";
 RL Biochemistry 27:7131-7140(1988).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YBR, AND C3H/HEJ, TISSUE=Sperm;
 RX MEDLINE=89384641; PubMed=2779578;
 RA Wawrzyniak C.J., Gallagher P.M., D'Amore M.A., Carter J.E.,
 RA Lund S.D., Rinchik E.M., Ganschow R.E.;
 RT "DNA determinants of structural and regulatory variation within the
 RT murine beta-glucuronidase gene complex.";
 RL Mol. Cell. Biol. 9:4074-4078(1989).
 [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88216590; PubMed=2835664;
 RA Funkenstein B., Leary S.L., Stein J.C., Catterall J.F.;
 RT "Genomic organization and sequence of the GUS-alpha allele of the
 RT murine beta-glucuronidase gene.";
 RL Mol. Cell. Biol. 8:1160-1168(1988).
 CC -1- FUNCTION: BETA-GLUCURONIDASE PLAYS AN IMPORTANT ROLE IN THE
 CC DEREGULATION OF DERMATAN AND KERATAN SULFATES.
 CC -1- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O = AN
 CC ALCOHOL + D-GLUCURONATE.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: J03047; AAA37696.1; -
 DR EMBL: J02836; AAA98623.1; -
 DR EMBL: M63836; AAA63309.1; -
 DR EMBL: M28540; AAA63307.1; -

DR EMBL: M28541; AAA63308.1; -
 DR EMBL: M19279; AAA37697.1; -
 DR PIR: A28954; A28954.
 DR PIR: A29977; A29977.
 DR HSSP: P08236; 1BHG.
 DR MGD: MGI:95874; Gus-s.
 DR Interpro: IPR001649; -
 DR Pfam: PF00703; Glyco_hydro_2; 1.
 DR PRINTS: PR00132; GLHYDRLASE2.
 DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
 DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
 KW Hydrolyase; Glycosidase; Lysosome; Glycoprotein; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 648
 FT ACT_SITE 447 447
 FT CARBOHYD 172 172
 FT CARBOHYD 416 416
 FT CARBOHYD 591 591
 FT CARBOHYD 627 627
 FT VARIANT 87 87
 FT VARIANT 233 233
 FT VARIANT 265 265
 FT VARIANT 320 320
 FT VARIANT 428 428
 FT VARIANT 616 616
 SQ SEQUENCE 648 AA; 74239 MW; 3DBCE5A5DB3B96D6 CRC64;

Query Match 54.4%; Score 37; DB 1; Length 648;
 Best Local Similarity 62.5%; Pred. No. 68;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 LEHOGGHV 12
 DB 140 VEHEGHL 147

RESULT 15
 BGIR_RAT
 ID BGLR_RAT STANDARD; PRT: 648 AA.
 AC P06760:
 DT 01-JAN-1988 (rel. 06, Created)
 DT 01-JAN-1988 (rel. 06, Last sequence update)
 DT 15-JUL-1998 (rel. 06, Last annotation update)
 DE BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31).
 GN GUSB OR GUS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=preputial gland;
 RX MEDLINE=87016933; PubMed=3463967;
 RA Nishimura Y., Rosenfeld M.G., Kreibich G., Gubler U., Sabatini D.D.,
 RA Adesnik M., Andy R.;
 RT "Nucleotide sequence of rat preputial gland beta-glucuronidase cDNA
 RT and in vitro insertion of its encoded polypeptide into microsomal
 RT membranes";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:7292-7296(1986).
 RN [2]
 RP SEQUENCE OF 14-648 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=88183378; PubMed=3355537;
 RA Powell P.P., Kyle J.W., Miller R.D., Pantano J., Grubb J.H., Sly W.S.;
 RT "Rat liver beta-glucuronidase. cDNA cloning, sequence comparisons and
 RT expression of a chimeric protein in COS cells.";
 RL Biochem. J. 250:547-555(1988).
 CC -1- FUNCTION: BETA-GLUCURONIDASE PLAYS AN IMPORTANT ROLE IN THE
 CC DEGRADATION OF DERMATAN AND KERATAN SULFATES.
 CC -1- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O = AN
 CC ALCOHOL + D-GLUCURONATE.
 CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.
 CC -1- PTM: UNDERGOES A POSTTRANSCRIPTIONAL PROTEOLYTIC CLEAVAGE NEAR ITS
 CC C-TERMINAL END, WHICH REDUCES ITS SIZE BY APPROXIMATELY 3 KDA. THE
 CC SITE OF THIS CLEAVAGE HAS AS YET NOT BEEN DETERMINED.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M13962; AAA41228.1; -
 DR EMBL: Y00717; CA68705.1; -
 DR PIR: A25047; A25047.
 DR PIR: S00345; S00345.
 DR HSSP: P08236; 1BHG.
 DR Interpro: IPR001649; -
 DR Pfam: PF00703; Glyco_hydro_2; 1.
 DR PRINTS: PR00132; GLHYDRLASE2.
 DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
 DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
 KW Hydrolyase; Glycosidase; Lysosome; Glycoprotein; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 648
 FT ACT_SITE 447 447
 FT CARBOHYD 172 172
 FT CARBOHYD 416 416
 FT CARBOHYD 591 591
 FT CARBOHYD 627 627
 FT CONFLICT 14 14
 FT CONFLICT 21 21
 FT CONFLICT 487 487
 SQ SEQUENCE 648 AA; 74793 MW; 5ADE8F5234F0907E CRC64;

Query Match 54.4%; Score 37; DB 1; Length 648;
 Best Local Similarity 62.5%; Pred. No. 68;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 LEHOGGHV 12
 DB 140 VEHEGHL 147

Search completed: July 6, 2001, 09:26:37
 Job time: 968 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:25:52 ; Search time 118.42 Seconds
(without alignments)
13.407 Million cell updates/sec

Title: US-09-437-912-2

Perfect score: 68

Sequence: 1 LDDDLHGGGHV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	66.2	236	10	081250
2	45	66.2	239	10	081253
3	45	66.2	239	10	09SBE1
4	44	64.7	541	2	0925A6
5	43	63.2	174	2	09KFE6
6	43	63.2	683	9	09T215
7	42	61.8	648	5	09VXP2
8	42	61.8	718	5	09XYV9
9	42	61.8	3282	2	09PEY9
10	42	61.8	3442	2	09PBE8
11	42	61.8	3455	5	09P906
12	40.5	59.6	536	5	045994
13	40	58.8	268	10	09LWC2
14	40	58.8	1025	2	054507
15	40	58.8	1025	2	P72532
16	40	58.8	1026	2	09X3R6
17	40	58.8	1046	2	084941
18	40	58.8	1715	6	09GLM4
19	40	58.8	1735	4	09HBL0

20	39	57.4	26	8	09T2S3	09T2S3 solanum tub
21	39	57.4	325	2	091AK5	091AK5 pseudomonas
22	39	57.4	515	8	096182	096182 wolinnella s
23	39	57.4	534	8	09T2S9	09T2S9 solanum tub
24	39	57.4	534	10	041445	041445 solanum tub
25	38.5	56.6	250	2	09FDL3	09FDL3 zymomonas m
26	38	55.9	81	2	09JVP9	09JVP9 neisseria m
27	38	55.9	122	10	003863	003863 zea mays (m
28	38	55.9	233	2	09KJX2	09KJX2 myxococcus
29	38	55.9	236	10	081254	081254 zea mays su
30	38	55.9	236	10	09SBE2	09SBE2 zea mays su
31	38	55.9	238	10	081249	081249 zea mays su
32	38	55.9	238	10	081255	081255 zea mays su
33	38	55.9	238	10	081257	081257 zea luxuria
34	38	55.9	238	10	081258	081258 zea luxuria
35	38	55.9	238	10	09SBE8	09SBE8 zea mays su
36	38	55.9	238	10	09SBE5	09SBE5 zea luxuria
37	38	55.9	238	10	09S626	09S626 zea luxuria
38	38	55.9	240	10	081252	081252 zea mays su
39	38	55.9	240	10	081256	081256 zea mays su
40	38	55.9	240	10	09SBE0	09SBE0 zea mays su
41	38	55.9	240	10	09SBE9	09SBE9 zea mays su
42	38	55.9	240	10	09SBE7	09SBE7 zea mays su
43	38	55.9	242	10	081251	081251 zea mays su
44	38	55.9	242	10	09SBE6	09SBE6 zea mays su
45	38	55.9	407	10	041750	041750 zea mays (m

ALIGNMENTS

-RESULT 1

ID 081250 PRELIMINARY; PRT; 236 AA.

AC 081250;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE GLOBULIN-1 (FRAGMENT).

OS Zea mays subsp. mays (maize).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;

OC Andropogoneae; Zea.

OX NCBI_TaxID=4578;

RN [1]

RP SEQUENCE FROM N.A.

RA Hilton H., Gaut B.S.;

RT "Speciation and domestication in maize and its wild relatives:

RT evidence from the globulin-1 gene."

RL Genetics 0:0-0(1998).

DR EMBL; AF064213; AAC31456.1; -

DR HSSP; P50477; ICAU.

DR Mendel; 31892; Zeama;1188;31892.

DR InterPro; IPR000901; -

DR InterPro; IPR001113; -

DR Pfam; PF00546; Seedstore_7s; 1.

DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.

FT NON_TER

FT 236

SQ SEQUENCE 236 AA; 27019 MW; 1F3D9BD92C032E05 CRC64;

Query Match 66.2%; Score 45; DB 10; Length 236;
Best Local Similarity 70.0%; Pred. NO. 4.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHGGGH 11
||| |||
Db 26 DDNLHNGGH 35

RESULT 2
081253 PRELIMINARY; PRT; 239 AA.

AC 081253;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE GLOBULIN-1 (FRAGMENT).
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4578;
RN [1]
RP SEQUENCE FROM N.A.
RA Hilton H., Gaut B.S.;
RT "Speciation and domestication in maize and its wild relatives:
RT evidence from the Globulin-1 gene.";
RL Genetics 0:0-0(1998).
DR EMBL; AF064216; AAC31459.1; -.
DR HSSP; P50477; ICAU.
DR Mendel; 31895; Zeama; 1188; 31895.
DR InterPro; IPR000901; -.
DR InterPro; IPR001113; -.
DR Pfam; PF00546; Seedstore_7s; 1.
DR PROSITE; PS00867; CPASASE_2; UNKNOWN_1.
FT NON_TER 239
SQ SEQUENCE 239 AA; 27384 MW; 628924A8D7BA7773 CRC64;

Query Match 66.2%; Score 45; DB 10; Length 239;
Best Local Similarity 70.0%; Pred. No. 4.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 DDDLEHOGGH 11
DB 26 DDNLHHGGH 35

RESULT 3
ID Q9SBE1 PRELIMINARY; PRT; 239 AA.
AC Q9SBE1;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE GLOBULIN-1 (FRAGMENT).
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4578;
RN [1]
RP SEQUENCE FROM N.A.
RA Hilton H., Gaut B.S.;
RT "Speciation and domestication in maize and its wild relatives:
RT evidence from the Globulin-1 gene.";
RL Genetics 0:0-0(1998).
DR EMBL; AF064218; AAC31461.1; -.
DR HSSP; P50477; ICAU.
DR InterPro; IPR000901; -.
DR InterPro; IPR001113; -.
DR Pfam; PF00546; Seedstore_7s; 1.
DR PROSITE; PS00867; CPASASE_2; UNKNOWN_1.
FT NON_TER 239
SQ SEQUENCE 239 AA; 27499 MW; 147C4F61F65307FA CRC64;

Query Match 66.2%; Score 45; DB 10; Length 239;
Best Local Similarity 70.0%; Pred. No. 4.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 DDDLEHOGGH 11
DB 26 DDNLHHGGH 35

RESULT 4
ID Q9ZSA6 PRELIMINARY; PRT; 541 AA.
AC Q9ZSA6;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE PUTATIVE LONG-CHAIN-FATTY-ACID COA LIGASE.
GN SC265.17.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Oliver K., Harris D.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE-97000351; PubMed-8843436;
RA Redenbach M., Kleser H.M., Denapalte D., Elchner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035478; CAB36604.1; -.
DR HSSP; P08659; ILCI.
DR InterPro; IPR000873; -.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KM ligase.
SQ SEQUENCE 541 AA; 59034 MW; CEB7374431F28CE5 CRC64;

Query Match 64.7%; Score 44; DB 2; Length 541;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 LDDLEHOGGHV 12
DB 349 MDDLEHRTGV 360

RESULT 5
ID Q9KTH6 PRELIMINARY; PRT; 174 AA.
AC Q9KTH6;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE BH0503 PROTEIN.
GN BH0503.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=8665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RA Takami H., Nakasone K., Takaki Y.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001308; BAB04222.1; -.
DR InterPro; IPR000182; -.
DR Pfam; PF00583; Acetyltransf; 1.
SQ SEQUENCE 174 AA; 19690 MW; 11D88B81C644F5E3 CRC64;

```

Query Match          63.2%; Score 43; DB 2; Length 174;
Best Local Similarity 58.3%; Pred. No. 7.1;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 DDDLEHOGGH 12
    ||| |||
DB 88 LNDMLHRCGHI 99

RESULT 6
Q9T215 PRELIMINARY; PRT; 683 AA.
AC Q9T215;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE REPRESSOR.
GN C.
OS Bacteriophage phi-C31.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10719;
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-NORMICH STOCK;
RA Smith M.C.M.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-NORMICH STOCK;
RA Hendrix R.W., Smith M.C.M., Burns N., Ford M.E., Hatfull G.F.;
RT "All the world's a phage.";
RT Proc. Natl. Acad. Sci. U.S.A. 96:2192-2197(1999).
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-NORMICH STOCK;
RA MEDLINE=99238410; Pubmed=10219087;
RA Smith M.C.M., Burns N., Wilson R.N., Gregory M.A.;
RT "The complete genome sequence of the Streptomyces temperate phage C31:
RT evolutionary relationships to other viruses.";
RT Nucleic Acids Res. 27:2145-2155(1999).
DR EMBL: AJ006589; CAA07123.1; -
SQ SEQUENCE 683 AA; 73980 MW; EC11A061EBCA4BD CRC64;

Query Match          63.2%; Score 43; DB 9; Length 683;
Best Local Similarity 70.0%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGGH 11
    ||| |||
DB 474 DDDVEROGAH 483

RESULT 7
Q9VXP2 PRELIMINARY; PRT; 648 AA.
AC Q9VXP2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE RHP PROTEIN.
GN RHP OR CG8497.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN RN
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chape M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bruttler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Galhart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003500; AAF48516.1; -
DR FlyBase: FBgn0026374; Rhp.
DR InterPro: IPR001478; -
DR Pfam: PF00595; PDZ; 1.
DR SMART: SM00228; PDZ; 1.
SQ SEQUENCE 648 AA; 72858 MW; 85EE4EBC6947D45D CRC64;

Query Match          61.8%; Score 42; DB 5; Length 648;
Best Local Similarity 70.0%; Pred. No. 42;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGGH 11
    ||| |||
DB 489 DDDDEHOGGY 498

RESULT 8
Q9XYV9 PRELIMINARY; PRT; 718 AA.
AC Q9XYV9;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE RHOPHILIN.
GN RHP OR CG8497.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN RN
RP SEQUENCE FROM N.A.

```

RA Edwards K.A., Kaneshiro K., Yamamoto D.;
 RT "Mutations in the Drosophila Rhoophilin gene at 13E.";
 RL Submitted (FEB-1999) to the EMBL/Genbank/DDbJ databases.
 DR EMBL: AF137025; AAD31273.1; -
 DR FlyBase; FBgn0026374; Rnp.
 DR InterPro; IPR000861; -
 DR InterPro; IPR001478; -
 DR Pfam; PF00595; PDZ. 1.
 DR Pfam; PF02185; HRI. 1.
 DR SMART; SM00228; PDZ. 1.
 SQ SEQUENCE 718 AA; 80826 MW; AF9DDOC57132AA31 CRC64;

Query Match 61.8%; Score 42; DB 5; Length 718;
 Best Local Similarity 70.0%; Pred. No. 46;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGGH 11
 DB 559 DDDEHGGY 568

RESULT 9
 ID Q9PEY9 PRELIMINARY; PRT; 3282 AA.
 AC Q9PEY9;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, last annotation update)
 DE HEMAGGLUTININ-LIKE SECRETED PROTEIN.
 GN XF0889.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xylella.
 NX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 MEDLINE-20365717; PubMed-10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
 Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 Facinoni A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vitorello C.B.,
 Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
 Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
 Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 DR EMBL; AE003928; AAF83699.1; -
 DR InterPro; IPR000267; -
 DR InterPro; IPR001424; -
 DR PROSITE; PS00144; ASN_GLN_ASE_1; UNKNOWN_1.
 DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN_1.
 SQ SEQUENCE 3282 AA; 343744 MW; F0796430BC9A2194 CRC64;

Query Match 61.8%; Score 42; DB 2; Length 3282;
 Best Local Similarity 41.7%; Pred. No. 2.2e+02;
 Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDDEHOGGH 12
 DB 472 LENDIDNRGHI 483

RESULT 10
 ID Q9PEB8 PRELIMINARY; PRT; 3442 AA.
 AC Q9PEB8;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, last annotation update)
 DE HEMAGGLUTININ-LIKE SECRETED PROTEIN.
 GN XF2196.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xylella.
 NX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 MEDLINE-20365717; PubMed-10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
 Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 Facinoni A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vitorello C.B.,
 Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
 Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
 Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 DR EMBL; AE004032; AAF84995.1; -
 DR InterPro; IPR000267; -
 DR InterPro; IPR001424; -
 DR PROSITE; PS00144; ASN_GLN_ASE_1; UNKNOWN_1.
 DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN_1.
 SQ SEQUENCE 3442 AA; 360148 MW; AAE30CDE923E3C6E CRC64;

Query Match 61.8%; Score 42; DB 2; Length 3442;
 Best Local Similarity 41.7%; Pred. No. 2.4e+02;
 Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDDEHOGGH 12
 DB 472 LENDIDNRGHI 483

```

RESULT 11
O9P906 PRELIMINARY; PRT: 3455 AA.
AC O9P906;
DT 01-OCT-2000 (TREMBLrel. 15 Created)
DE 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HEMAGGLUTININ-LIKE SECRETED PROTEIN.
GN XF2775.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Bala J.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facinunzi A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Graber A.,
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.V., Madela A.M.B.N., Madela H.M.F., Marinho C.L.,
RA Marques M.V., Miracca E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.A.L., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza M.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchino M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zaço M.A., Zatz M., Zeldanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
DR EMBL: AE004082; AAF55560.1;
DR InterPro: IPR000267;
DR InterPro: IPR001424;
DR PROSITE: PS00144; ASN_GLM_ASE_1; UNKNOWN_1.
DR PROSITE: PS00087; SOD_CU_ZN_1; UNKNOWN_1.
SQ SEQUENCE 3455 AA; 360947 MW; 4CD692CB8752FDAA CRC64;

Query Match 61.8%; Score 42; DB 2; Length 3455;
Best Local Similarity 41.7%; Pred. No. 2.4e+02;
Matches 5; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDDDLHOGGHV 12
Db 472 LENDIDNRGHI 483

RESULT 12
O45994 PRELIMINARY; PRT: 536 AA.
AC O45994;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE ZK1053.2 PROTEIN.
GN ZK1053.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```

```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Smalton N., Smith A., Sonnenhammer E., Staden K., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston J.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL: Z82084; CAB04976.1;
SQ SEQUENCE 536 AA; 59559 MW; 01F70E269ACFD7C4 CRC64;

Query Match 59.6%; Score 40.5; DB 5; Length 536;
Best Local Similarity 53.3%; Pred. No. 61;
Matches 8; Conservative 4; Mismatches 0; Indels 3; Gaps 1;

OY 1 LDDDLHOGGHV 12
Db 511 LDDVDHMRPHGGV 525

RESULT 13
O9LWC2 PRELIMINARY; PRT: 268 AA.
AC O9LWC2;
DT 01-OCT-2000 (TREMBLrel. 15 Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaeae;
OC Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV_NIPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0483F08."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002094; BAA96222.1;
SQ SEQUENCE 268 AA; 28421 MW; 2A5A4BA7DD8B84F4 CRC64;

Query Match 58.8%; Score 40; DB 10; Length 268;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 DDDLEHOGGH 11
Db 169 EDEVESOGGH 178

RESULT 14
O54507 PRELIMINARY; PRT: 1025 AA.
AC O54507;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

```

Job time: 988 sec

DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE SERUM OPACITY FACTOR.
 GN SOF.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxId=1314;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=M2;
 RX MEDLINE=95122198; PubMed=7822031;
 RA Rakonjac J.V., Robbins J.C., Fischetti V.A.;
 RT "DNA sequence of the serum opacity factor of group A streptococci:
 RT Identification of a fibronectin-binding repeat domain.";
 RL Infect. Immun. 63:622-631(1995).
 DR EMBL: U02290; AA85219.1; -.
 DR HSP; P02188; 1HRM.
 DR InterPro: IPR001899; -.
 DR InterPro: IPR002035; -.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 SQ SEQUENCE 1025 AA; 112770 MW; 327C9C1DFBBD0394 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 1025;
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGGH 11
 1:|||||
 DB 809 DEDLEISGCH 818

RESULT 15
 ID P72532 PRELIMINARY; PRT; 1025 AA.
 AC P72532;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE FIBRONECTIN-BINDING PROTEIN II.
 GN SFBII.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxId=1314;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLINICAL ISOLATE A75;
 RX MEDLINE=96020668; PubMed=7476200;
 RA Kreikemeyer B., Talay S.R., Chatwal G.S.;
 RT "Characterization of a novel fibronectin-binding surface protein in
 RT group A streptococci.";
 RL Mol. Microbiol. 17:137-145(1995).
 DR EMBL: X83303; CAA58282.1; -.
 DR HSP; P02188; 1HRM.
 DR InterPro: IPR001899; -.
 DR InterPro: IPR002035; -.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 SQ SEQUENCE 1025 AA; 112680 MW; CE6E52FE45CFE5D5 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 1025;
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DDDLEHOGGH 11
 1:|||||
 DB 809 DEDLEISGCH 818

Search completed: July 6, 2001, 09:25:52

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:20 ; Search time 56.74 Seconds
(without alignments)
4.260 Million cell updates/sec

Title: US-09-437-912-3
Perfect score: 79
Sequence: 1 GHRKHGHCNGK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	64.6	1213	1 US-08-188-582-20	Sequence 20, Appl
2	51	64.6	1213	1 US-08-646-715-20	Sequence 20, Appl
3	50	63.3	113	2 US-08-918-727-7	Sequence 7, Appl
4	50	63.3	113	2 US-09-205-680A-7	Sequence 7, Appl
5	50	63.3	345	2 US-08-758-621-14	Sequence 14, Appl
6	50	63.3	345	4 US-09-107-858-14	Sequence 14, Appl
7	48	60.8	339	2 US-08-758-621-2	Sequence 2, Appl
8	48	60.8	339	4 US-09-107-858-2	Sequence 2, Appl
9	48	60.8	575	2 US-09-032-315-8	Sequence 8, Appl
10	48	60.8	575	2 US-08-993-318A-8	Sequence 8, Appl
11	48	60.8	575	4 US-09-399-886-8	Sequence 8, Appl
12	48	60.8	575	4 US-09-396-260-8	Sequence 8, Appl
13	48	60.8	576	1 US-08-172-331B-2	Sequence 26, Appl
14	47	59.5	617	1 US-08-137-614A-26	Sequence 26, Appl
15	47	59.5	637	3 US-08-072-064-1	Sequence 1, Appl
16	47	59.5	637	3 US-08-072-064-4	Sequence 4, Appl
17	47	59.5	637	3 US-08-072-064-6	Sequence 6, Appl
18	47	59.5	637	3 US-08-072-064-8	Sequence 8, Appl
19	47	59.5	637	5 PCT-US92-08558-1	Sequence 1, Appl
20	46	58.2	1199	4 US-09-208-742-2	Sequence 2, Appl
21	45	57.0	493	1 US-08-362-512A-4	Sequence 4, Appl
22	44	55.7	515	4 US-08-942-012B-32	Sequence 32, Appl
23	43	54.4	18	1 US-08-293-284A-64	Sequence 64, Appl
24	43	54.4	18	1 US-08-346-849-64	Sequence 64, Appl
25	43	54.4	313	3 US-08-686-528A-3	Sequence 3, Appl
26	43	54.4	313	4 US-09-456-287-3	Sequence 3, Appl
27	43	54.4	337	3 US-08-686-528A-2	Sequence 2, Appl

28	43	54.4	337	4 US-09-456-287-2	Sequence 2, Appl
29	43	54.4	802	2 US-08-007-107-4	Sequence 4, Appl
30	41.5	52.5	355	2 US-08-758-621-4	Sequence 4, Appl
31	41.5	52.5	355	4 US-09-107-858-4	Sequence 4, Appl
32	41	51.9	763	2 US-08-677-862-2	Sequence 2, Appl
33	41	51.9	763	2 US-09-252-571-2	Sequence 2, Appl
34	41	51.9	763	3 US-09-434-065-2	Sequence 8, Appl
35	40	50.6	16	1 US-08-346-849-53	Sequence 53, Appl
36	40	50.6	16	2 US-08-293-284A-53	Sequence 53, Appl
37	40	50.6	21	2 US-08-651-818A-21	Sequence 21, Appl
38	40	50.6	23	1 US-08-480-604A-24	Sequence 24, Appl
39	40	50.6	23	3 US-08-405-496A-24	Sequence 8, Appl
40	40	50.6	24	3 US-08-584-031-8	Sequence 8, Appl
41	40	50.6	24	3 US-08-780-496-8	Sequence 19, Appl
42	40	50.6	53	2 US-08-651-818A-19	Sequence 23, Appl
43	40	50.6	54	2 US-08-651-818A-23	Sequence 23, Appl
44	40	50.6	60	1 US-08-255-457-1	Sequence 1, Appl
45	40	50.6	60	2 US-09-115-032-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT 1
US-08-188-582-20
: Sequence 20, Application US/08188582
: Patent No. 5534410
:
: GENERAL INFORMATION:
: APPLICANT: Tjian, Robert
: APPLICANT: Comai, Lucio
: APPLICANT: Dynlacht, Brian D.
: APPLICANT: Hoey, Timothy
: APPLICANT: Ruppert, Siegfried
: APPLICANT: Tanese, Naoko
: APPLICANT: Weinzierl, Robert O.J.
: TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
: TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TFs AND METHODS OF USE
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/188,582
: FILING DATE: 28-JAN-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-57650-2/AT/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELEX: 910 272299
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1213 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-188-582-20
:
: Query Match 64.6%, Score 51, DB 1, Length 1213;

```

Best Local Similarity 72.7%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HKHKGHGCHK 12
|||||

Db 1160 HKHKGHGCHK 1170

RESULT 2

US-08-646-715-20

Sequence 20, Application US/08646715
Patent No. 5637686

GENERAL INFORMATION:

APPLICANT: Tjian, Robert

APPLICANT: Comai, Lucio

APPLICANT: Dynlacht, Brian D.

APPLICANT: Hoey, Timothy

APPLICANT: Ruppert, Siegfried

APPLICANT: Tanese, Naoko

APPLICANT: Wang, Edith

APPLICANT: Weinzierl, Robert O.J.

TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS

TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/646,715

FILING DATE: 09-MAY-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/188,582

FILING DATE: 28-JAN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 1213 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-646-715-20

Query Match 64.6%; Score 51; DB 1; Length 1213;

Best Local Similarity 72.7%; Pred. No. 6.7;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HKHKGHGCHK 12
|||||

Db 1160 HKHKGHGCHK 1170

RESULT 3

US-08-918-727-7

Sequence 7, Application US/08918727

Patent No. 5849528

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Lal, Preeti

APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN S100 PROTEINS

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/918,727

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0373 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 113 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 488157

US-08-918-727-7

Query Match 63.3%; Score 50; DB 2; Length 113;

Best Local Similarity 66.7%; Pred. No. 0.91;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GKHKHGCHK 12
|||

Db 102 GHDHRGKCGK 113

RESULT 4

US-09-205-680A-7

Sequence 7, Application US/09205680A

Patent No. 6103497

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

APPLICANT: Lal, Preeti

APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN S100 PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/205,680A
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Colette C. Muenzen
REGISTRATION NUMBER: 39,784
REFERENCE/DOCKET NUMBER: PF-0373 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 488157
US-09-205-680A-7

Query Match 63.3%; Score 50; DB 3; Length 113;
Best Local Similarity 66.7%; Pred. NO. 0.91;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GHKHHGHGCK 12
DB 102 GHDHGHGCK 113

RESULT 5
US-08-758-621-14
Sequence 14, Application US/08758621
Patent No. 5846821
GENERAL INFORMATION:
APPLICANT: Guerino, Mary Lou, and Elde, David J.
TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,621
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,578
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Silveri, Jean M.
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: DCI-099CP

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-758-621-14

Query Match 63.3%; Score 50; DB 2; Length 345;
Best Local Similarity 77.8%; Pred. NO. 2.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GHKHHGHG 9
DB 167 GHHSHGHG 175
RESULT 6
US-09-107-858-14
Sequence 14, Application US/09107858
Patent No. 6162900
GENERAL INFORMATION:
APPLICANT: Guerino, Mary Lou et al.
TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR
FILE REFERENCE: DCI-099CPDV
CURRENT APPLICATION NUMBER: US/09/107,858
CURRENT FILING DATE: 1998-06-30
EARLIER APPLICATION NUMBER: 08/758,621
EARLIER FILING DATE: 1996-11-27
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 345
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-107-858-14

Query Match 63.3%; Score 50; DB 4; Length 345;
Best Local Similarity 77.8%; Pred. NO. 2.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GHKHHGHG 9
DB 167 GHHSHGHG 175

RESULT 7
US-08-758-621-2
Sequence 2, Application US/08758621
Patent No. 5846821
GENERAL INFORMATION:
APPLICANT: Guerino, Mary Lou, and Elde, David J.
TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,621

FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,578
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Silveri, Jean M.
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: DC1-099CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 339 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-758-621-2

Query Match 60.8%; Score 48; DB 2; Length 339;
Best Local Similarity 87.5%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 HKHGHG 11
11111111
Db 154 HGHGHG 161

RESULT 8
US-09-107-858-2
Sequence 2, Application US/09107858
Patent No. 6162900
GENERAL INFORMATION:
APPLICANT: Guerinet, Mary Lou et al.
FILE REFERENCE: DC1-099CPDV
CURRENT APPLICATION NUMBER: US/09/107,858
CURRENT FILING DATE: 1998-06-30
EARLIER APPLICATION NUMBER: 08/758,621
EARLIER FILING DATE: 1996-11-27
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 339
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-107-858-2

Query Match 60.8%; Score 48; DB 4; Length 339;
Best Local Similarity 87.5%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 HKHGHG 11
11111111
Db 154 HGHGHG 161

RESULT 9
US-09-032-315-8
Sequence 8, Application US/09032315
Patent No. 5985818
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5985818 of No. 5985818th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 5200,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-315-8

Query Match 60.8%; Score 48; DB 2; Length 575;
Best Local Similarity 64.3%; Pred. No. 8.4;
Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

OY 2 HKHKGH---GHG 11
1111111111
Db 331 HKHKGHGLSGHG 344

RESULT 10
US-08-993-318A-8
Sequence 8, Application US/08993318A
Patent No. 5998353
GENERAL INFORMATION:
APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Grethe
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5998353 of No. 5998353th America
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,318A
FILING DATE: December 18, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-318A-8

Query Match 60.8%; Score 48; DB 2; Length 575;
Best Local Similarity 64.3%; Pred. No. 8.4;
Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

OY 2 HKHKGH---GHG 11
|||||
DB 331 HKHKGRLSGHG 344

RESULT 11
US-09-399-886-8
Sequence 8, Application US/09399886
Patent No. 6140092
GENERAL INFORMATION:
APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Grethe
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 61400920 No. 6140092disk of No. 6140092th America
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/399,886
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,318
FILING DATE: December 18, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-399-886-8

Query Match 60.8%; Score 48; DB 4; Length 575;
Best Local Similarity 64.3%; Pred. No. 8.4;
Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
OY 2 HKHKGH---GHG 11
|||||
DB 331 HKHKGRLSGHG 344

RESULT 12

US-09-396-260-8
Sequence 8, Application US/09396260
Patent No. 6184015
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan

APPLICANT: Xu, Feng
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 61840150 No. 6184015disk of No. 6184015th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/396,260
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 5200.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-396-260-8

Query Match 60.8%; Score 48; DB 4; Length 575;
Best Local Similarity 64.3%; Pred. No. 8.4;
Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

OY 2 HKHKGH---GHG 11
|||||
DB 331 HKHKGRLSGHG 344

RESULT 13
US-08-172-331B-2
Sequence 2, Application US/08172331B
Patent No. 5480801
GENERAL INFORMATION:
APPLICANT: Mahleithner, Jill A.
APPLICANT: Christensen, Bjorn E.
APPLICANT: Schneider, Palle
TITLE OF INVENTION: PURIFIED PH NEUTRAL LACCASES AND NUCLEIC
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 54808010 No. 5480801disk of No. 5480801th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA

ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/172,331B
FILING DATE: 22-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,230
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,827
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/162,827
FILING DATE: 03-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lowmey Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4052.020-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-172-331B-2

Query Match 60.8%; Score 48; DB 1; Length 576;
Best Local Similarity 64.3%; Pred. No. 8.5;
Matches 9; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

OY 2 HKHKGH---GHG 11
|||||
DB 332 HKHKGRLSGHG 345

RESULT 14
US-08-137-614A-26
Sequence 26, Application US/08137614A
Patent No. 5487976
GENERAL INFORMATION:
APPLICANT: Soderlund, David M.
APPLICANT: Knipfle, Douglas C.
APPLICANT: Henderson, Joseph E.
TITLE OF INVENTION: Gene Encoding An Insect
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,614A
FILING DATE: 15-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timlan, Susan J.

REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1636
TELEFAX: (716)263-1600
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-137-614A-26

Query Match 59.5%; Score 47; DB 1; Length 617;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGH 10
|:| |||||
DB 434 GPEHGHHGH 443

RESULT 15
US-08-072-064-1
Sequence 1, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FERRENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,064
FILING DATE: 19930602
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 770,881
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OPHD-00574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster
POSITION IN GENOME:
CHROMOSOME/SEGMENT: III; polytene subregion 66F
MAP POSITION: approximately map unit 26
US-08-072-064-1

Query Match 59.5%; Score 47; DB 3; Length 637;

Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHHGHGH 10
| : | | | | |
Db 455 GPEHHGHGH 464

Search completed: July 6, 2001, 09:10:21
Job time: 187 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:17:58 ; Search time 73.59 Seconds
(without alignments)
12.421 Million cell updates/sec

Title: US-09-437-912-3

Perfect score: 79

Sequence: 1 GHKHKHGHHGK 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	644	1 KGHU1	kininogen, HMW pre
2	72	91.1	619	1 KGBH2	kininogen, HMW II
3	72	91.1	621	1 KGBH1	kininogen, HMW I P
4	62	78.5	189	2 C81428	peptidyl-prolyl ci
5	62	78.5	389	2 B96635	hypothetical prote
6	62	78.5	398	2 T02681	probable zinc tran
7	62	78.5	490	2 T36920	hypothetical prote
8	62	78.5	503	2 S54302	zinc transporter 2
9	61	77.2	507	2 S54303	zinc transporter pro
10	61	77.2	286	2 S07193	chorion protein s3
11	60	75.9	18	2 B32473	histidine-rich pro
12	60	75.9	670	2 F36791	hypothetical prote
13	60	75.9	2038	2 A43742	female sterile hom
14	59	74.7	535	2 S66148	gene pipsqueak pro
15	59	74.7	1085	2 S66149	hypothetical prote
16	58	73.4	318	2 T23089	related to spliced
17	57	72.2	515	2 T49714	K-kininogen, HMW P
18	56	70.9	264	2 C25486	K-kininogen, HMW P
19	56	70.9	639	2 A25486	kininogen, HMW I P
20	55	69.6	224	2 T34937	hypothetical prote
21	55	69.6	336	1 S75947	hypothetical prote
22	55	69.6	378	2 T49164	zinc transporter-1
23	54.5	69.0	290	2 C27115	K-kininogen, LMW P
24	54.5	69.0	315	2 A27115	major acute phase
25	54	68.4	199	2 T48059	hypothetical prote
26	54	68.4	251	2 T34168	hypothetical prote
27	54	68.4	334	2 T47986	zinc transporter-1
28	54	68.4	344	2 T05310	hypothetical prote
29	54	68.4	1221	2 T13283	probable transcrip

30	54	68.4	1891	2 T13594	hypothetical prote
31	54	68.4	1920	2 T13893	gene hindsight pro
32	53.5	67.7	173	2 T51469	glycine/proline-ri
33	53	67.1	160	2 T07180	hypothetical prote
34	53	67.1	177	2 S65780	glycine/proline-ri
35	53	67.1	196	2 D85999	hypothetical prote
36	53	67.1	196	2 A49987	probable fkbp-type
37	53	67.1	529	2 T08684	hypothetical prote
38	52	65.8	110	2 T07618	cold stress protei
39	52	65.8	314	2 T35241	hypothetical prote
40	52	65.8	436	2 T36240	MHC H-2K/T-W5-link
41	51	64.6	203	2 T39079	hypothetical prote
42	51	64.6	457	2 S39079	puff C-8 protein -
43	51	64.6	1028	2 A56038	DNA-binding protei
44	51	64.6	1213	2 S16356	ovo protein - trul
45	51	64.6	1213	2 A54063	TAA-binding prote

ALIGNMENTS

RESULT 1

KGHDH1

kininogen, HMW precursor [validated] - human

N:Alternate names: alpha-2-chiol proteinase inhibitor; preprokininogen; prokininogen

C:Contains: bradykinin (Kallidin I); HMW kininogen I; HMW kininogen II; low molecular

C:Species: Homo sapiens (man)

C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000

C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030;

R:Ohkubo, I.; Kuzachi, K.; Takasawa, T.; Shiohawa, H.; Sasaki, M.

Biochemistry 23, 5691-5697, 1984

A:Title: Isolation of a human cDNA for alpha-2-chiol proteinase inhibitor and its ide

A:Reference number: A90490; MUID:85122621

A:Accession: A01279

A:Molecule type: mRNA

A:Residues: 1-389 <OHK>

A:Cross-references: GB:M1437; NID:G177889

R:Takagaki, Y.; Kitamura, N.; Nakamishi, S.

J. Biol. Chem. 260, 8601-8609, 1985

A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and 1

A:Reference number: A92544; MUID:85234582

A:Accession: A25276

A:Molecule type: mRNA

A:Residues: 1-592, 'T', 594-644 <TAK>

A:Cross-references: GB:M1437; NID:G186751; PIDN:AAB59550.1; PID:G386852

R:Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.

FEBS Lett. 321, 93-97, 1993

A:Title: Cloning, expression and characterization of human kininogen domain 3.

A:Reference number: S32422; MUID:93223854

A:Accession: S32422

A:Molecule type: mRNA

A:Residues: 'ANSM', 253-377 <AUE>

A:Note: differences are due to known cloning artifacts

R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.

Eur. J. Biochem. 152, 307-314, 1995

A:Title: The amino acid sequence of the light chain of human high-molecular-mass kinin

A:Reference number: A91153; MUID:86030270

A:Accession: A91153

A:Molecule type: protein

A:Residues: 379-644 <LOT>

A:Note: the bradykinin sequence preceding the light chain sequence was not determined

R:Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.

Eur. J. Biochem. 154, 471-478, 1986

A:Title: Completion of the primary structure of human high-molecular-mass kininogen.

A:Reference number: A24871; MUID:86108361

A:Accession: A24871

A:Molecule type: protein

A:Residues: 'Z', 20-380 <KEU>

R:Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.

in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp. 85-89, Plenum Press, New

A:Title: Amino acid sequence of the light chain of human high molecular mass kininoge

A:Reference number: A27899

A:Accession: A27899

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEI2>
 R:Minidrou, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kindin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88209021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A1905; MUID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAE>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plas
 A:Reference number: A34030; MUID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
 A:Reference number: S02482; MUID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19; 189-192; 310-314; 381-389 <LENI>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Experimental source: urine
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55239; MUID:95251593
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Stracek, J.; Maachi, F.; Le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; MUID:96033974
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitaawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evolu
 A:Reference number: A92345; MUID:85234583
 A:Contents: annotation; gene organization

R: Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91453; MUID:90253622
 A:Contents: annotation; bradykinin
 A:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is i
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 X:Proline residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F:19-131/Domain: cystatin homology <CT1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19-Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status expert
 F:28-614,83-94,107-126,142-216,229-248,264-267,328-340,351-370/Disulfide bond
 F:48/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:165,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status
 F:577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 79; DB 1; Length 644;
 Best Local Similarity 100.0%; Pred. No. 0.00029;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GHKHHGHGCK 12
 Db 494 GHKHHGHGCK 505

RESULT 2
 KGB0H2
 N:kininogen, HMW II precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence; revision 14-Nov-1983 #text_change 22-Jun-1999
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 376-391 <KAT>
 R:Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.

J. Biochem. 79, 1201-1222, 1976
A:Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino a
A:Reference number: A91941; MUID:76260155
A:Accession: A91941
A:Molecule type: protein
A:Residues: 387-455 <HAN>
A:Note: 398-Pro, 401-Val, and 455-Lys were also found
R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Am
A:Reference number: A91938; MUID:75170265
A:Accession: A91938
A:Molecule type: protein
A:Residues: 456-496 <HA2>
R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwana
J. Biol. Chem. 262, 2768-2779, 1987
A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A:Reference number: A92627; MUID:87137530
A:Accession: B29559
A:Molecule type: protein
A:Residues: 72', 20-104, 'E', 106-256, 'XX', 257-376 <SUE>
R:Rottsch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A:Reference number: A91153; MUID:86030270
A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A:Title: Disulfide bonds in bovine HMW kininogens.
A:Reference number: A94300
A:Contents: annotation; disulfide bonds
A:Note: article in Japanese
C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-619/Product: HMW kininogen II #status predicted <MAT>
F:19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F:19-130/Domain: cystatin homology <CY1>
F:141-252/Domain: cystatin homology <CY2>
F:261-372/Domain: cystatin homology <CY3>
F:377-386/Product: bradykinin (kallidin II) #status experimental <KBDY>
F:378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
F:387-619/Product: HMW kininogen II light chain #status experimental <LCH>
F:418-468/Region: glycine/histidine/lysine-rich
F:119/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experim
F:27-589, 82-93, 106-125, 141-144, 205-217, 228-247, 261-264, 325-337, 348-367/Disulfide bonds:
F:47/Binding site: carbohydrate (Asn) (covalent) #status absent
F:87, 168, 169, 204, 280/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F:197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
F:380/Modified site: 4-hydroxyproline (Pro) #status predicted
F:386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:396, 400, 404, 510/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:397, 398, 518, 522, 534, 546, 551, 568/Binding site: carbohydrate (Thr) (covalent) #status ex
F:496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 91.1%; Score 72; DB 1; Length 619;
Best Local Similarity 91.7%; Pred. NO. 0.0029;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Cy 1 GKKRKHGCHGK 12
DB 470 GKKRKHGCHGK 481
RESULT 3

KGB0H1
kininogen, HMW I precursor - bovine
N:Alternate names: alpha-2-ubiquitin-proteinase inhibitor; prekallikrein
N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C:Accession: A01281; A91923; A91938; A29559
R:Kikamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A:Title: A single gene for bovine high molecular weight and low molecular weight kin
A:Reference number: A93317; MUID:84014106
A:Accession: A01281
A:Molecule type: mRNA
A:Residues: 1-621 <KIT>
A:Cross-references: GB:V01491; GB:R01757; NID:g491; PIDN:CAA24735.1; PID:g492
R:Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
A:Reference number: A91923; MUID:70180420
A:Accession: A91923
A:Molecule type: protein
A:Residues: 378-393 <KAT>
R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen.
A:Reference number: A91938; MUID:75170265
A:Accession: A91938
A:Molecule type: protein
A:Residues: 458-498 <HAN>
R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwana
J. Biol. Chem. 262, 2768-2779, 1987
A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions o
A:Reference number: A92627; MUID:87137530
A:Accession: A29559
A:Molecule type: protein
A:Residues: 72', 20-123, 'I', 129-378 <SUE>
R:Rottsch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kin
A:Reference number: A91153; MUID:86030270
A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A:Title: Disulfide bonds in bovine HMW kininogens.
A:Reference number: A94300
A:Contents: annotation; disulfide bonds
A:Note: article in Japanese
C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is i
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-621/Product: HMW prokininogen I #status predicted <MAT>
F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F:19-130/Domain: cystatin homology <CY1>
F:141-252/Domain: cystatin homology <CY2>
F:263-374/Domain: cystatin homology <CY3>
F:379-388/Product: bradykinin (kallidin II) #status experimental <KBDY>
F:380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F:417-468/Region: glycine/histidine/lysine-rich
F:119/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F:27-591, 82-93, 106-125, 141-144, 205-217, 228-247, 263-266, 327-339, 350-369/Disulfide bond
F:87, 168, 169, 204/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F:197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:376-379/Cleavage site: Met-Lys (kallikrein) #status predicted
F:382/Modified site: 4-hydroxyproline (Pro) #status experimental
F:388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:398, 406, 512/Binding site: carbohydrate (Ser) (covalent) #status experimental

Db 49 GHDRHGHGHC 59

RESULT 8

zinc transporter Znt-1 - mouse
S54302
C:Species: Mus musculus (house mouse)
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 05-Nov-1999
C:Accession: S54302
R:Palmiter, R.D.; Findley, S.D.
EMBO J. 14, 639-649, 1995
A:Title: Cloning and functional characterization of a mammalian zinc transporter that cc
A:Reference number: S54302; MUID:95188868
A:Accession: S54302
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <PAL>
A:Cross-references: EMBL:U17132; NID:9577840; PIDN:AAA79233.1; PID:9577841
C:Genetics:
A:Introns: 203/1

Query Match 78.5%; Score 62; DB 2; Length 503;
Best Local Similarity 81.8%; Pred. No. 0.066;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGHGHC 11
||| |||||
Db 145 GHGSHGHGHC 155

RESULT 9

zinc transport protein Znt-1 - rat
S54303
N:Alternate names: zinc transporter Znt-1
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 05-Nov-1999
C:Accession: S54303
R:Palmiter, R.D.; Findley, S.D.
EMBO J. 14, 639-649, 1995
A:Title: Cloning and functional characterization of a mammalian zinc transporter that cc
A:Reference number: S54302; MUID:95188868
A:Accession: S54303
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-507 <PAL>
A:Cross-references: EMBL:U17133; NID:9577842; PIDN:AAA79234.1; PID:9577843

Query Match 78.5%; Score 62; DB 2; Length 507;
Best Local Similarity 81.8%; Pred. No. 0.066;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGHGHC 11
||| |||||
Db 145 GHGSHGHGHC 155

RESULT 10

chorion protein s36 - fruit fly (Drosophila melanogaster)
S07193
C:Species: Drosophila melanogaster
C:Date: 01-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 21-Jul-2000
C:Accession: S07193
R:Spradling, A.C.; de Cicco, D.V.; Wakimoto, B.T.; Levine, J.F.; Kafayan, L.J.; Cooley, E.M.O. J. 6, 1045-1053, 1987
A:Title: Amplification of the X-linked Drosophila chorion gene cluster requires a region
A:Reference number: S07193; MUID:87246506
A:Accession: S07193
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <SPR>

A:Cross-references: EMBL:X05245; NID:97725; PIDN:CAA28870.1; PID:97726
C:Genetics:
A:Gene: FlyBase:Cb36
A:Cross-references: FlyBase:Fbgn0000359
A:Introns: 16/3

Query Match 77.2%; Score 61; DB 2; Length 286;
Best Local Similarity 75.0%; Pred. No. 0.054;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGHGHC 12
||| |||||
Db 27 GHGSHGHGHCQ 38

RESULT 11

histidine-rich protein C, peptide P-5 - liver fluke (fragment)
B32473
C:Species: Fasciola hepatica (liver fluke)
C:Date: 25-Sep-1989 #sequence_revision 03-May-1994 #text_change 26-May-2000
C:Accession: B32473
R:Walte, J.H.; Rice-Ficht, A.C.
Biochemistry 28, 6104-6110, 1989
A:Title: A histidine-rich protein from the vitellaria of the liver fluke Fasciola hep
A:Reference number: A32473; MUID:89375343
A:Accession: B32473
A:Molecule type: protein
A:Residues: 1-18 <MALD>
A:Note: 18-Gly and 18-His were also found
C:Superfamily: period clock protein; EGF homology
C:Keywords: egg yolk
F:1.5/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental

Query Match 75.9%; Score 60; DB 2; Length 18;
Best Local Similarity 81.8%; Pred. No. 0.0054;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGHGHC 11
||| |||||
Db 6 GHGSHGHGHC 16

RESULT 12

hypothetical protein ORF50 - ictalurid herpesvirus 1 (strain auburn 1)
F36791
C:Species: Ictalurid herpesvirus 1
A:Note: host Ictalurus punctatus (channel catfish)
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 31-Jan-2000
C:Accession: F36791
R:Davidson, A.J.
submitted to GenBank, January 1992
A:Description: Channel catfish virus: a new type of herpesvirus.
A:Reference number: A36804
A:Accession: F36791
A:Molecule type: DNA
A:Residues: 1-670 <DAV>
A:Cross-references: GB:M75136; NID:9331209; PIDN:AAA88153.1; PID:9331260
R:Davidson, A.J.
Virology 186, 9-14, 1992
A:Title: Channel catfish virus: a new type of herpesvirus.
A:Reference number: A39447; MUID:92087490
A:Contents: annotation
A:Note: neither protein nor nucleic acid sequence is given
C:Genetics:
A:Gene: 50
C:Superfamily: period clock protein; EGF homology

Query Match 75.9%; Score 60; DB 2; Length 670;
Best Local Similarity 81.8%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GKKKHGHG 11
 11111111
 DB 636 GHGHGHGHG 646

RESULT 13

A43742 female sterile homeotic protein, 205K - fruit fly (Drosophila melanogaster)
 N:Alternate names: membrane protein fsh, 205K
 N:Contains: female sterile homeotic protein, 110K
 C:Species: Drosophila melanogaster
 C>Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 20-Sep-1999
 C:Accession: A43742; B43742
 R:Weber, U.; Siegel, V.; Mlodzik, M.
 EMO J. 14, 6247-6257, 1995
 A:Title: The Drosophila fsh locus, a maternal effect homeotic gene, encodes apparent mem
 A:Reference number: A43742; MUID:89276730
 A:Accession: A43742
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2038 <HAY>
 A:Cross-references: EMBL:M23221; NID:g157452; PIDN:AAA28540.1; PID:g157453
 A:Accession: B43742
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1106 <HAZ>
 A:Cross-references: EMBL:M23222
 C:Genetics:
 A:Gene: fsh
 A:Cross-references: FlyBase:FBgn0004656
 C:Superfamily: unassigned bromodomain proteins; bromodomain homology
 F:1-2038/Product: female sterile homeotic protein, 205K #status predicted <MA2>
 F:1-1106/Product: female sterile homeotic protein, 110K #status predicted <MAT>
 F:59-116/Domain: bromodomain homology <BRO1>
 F:503-560/Domain: bromodomain homology <BRO2>

Query Match 75.9%; Score 60; DB 2; Length 2038;
 Best Local Similarity 81.8%; Pred. No. 0.49;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GKKKHGHG 11
 11111111
 DB 597 GHGHGHGHG 607

RESULT 14

S66148 gene pipsqueak protein A short form - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000
 C:Accession: S66148
 R:Weber, U.; Siegel, V.; Mlodzik, M.
 EMO J. 14, 6247-6257, 1995
 A:Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for t
 A:Reference number: S66148; MUID:96134923
 A:Accession: S66148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-535 <WEB>
 A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62473.1; PID:g1149499
 C:Genetics:
 A:Gene: pipsqueak
 C:Superfamily: POZ domain homology
 F:21-123/Domain: POZ domain homology <POZ>

Query Match 74.7%; Score 59; DB 2; Length 535;
 Best Local Similarity 80.0%; Pred. No. 0.19;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 HKKHGHG 11
 11111111
 DB 334 HEHNGHGHG 343

RESULT 15

S66149 gene pipsqueak protein A long form - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 01-Dec-2000
 C:Accession: S66149; S66150; T45461
 R:Weber, U.; Siegel, V.; Mlodzik, M.
 EMO J. 14, 6247-6257, 1995
 A:Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up fo
 A:Reference number: S66148; MUID:96134923
 A:Accession: S66149
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1085 <WEB>
 A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62474.1; PID:g1149500
 A:Accession: S66150
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 'MQ', 428-1085 <WE2>
 A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62475.1; PID:g1149501
 R:Horowitz, H.; Berg, C.A.
 Development 122, 1859-1871, 1996
 A:Title: The Drosophila pipsqueak gene encodes a nuclear BTB-domain-containing protel
 A:Reference number: Z22972; MUID:96232300
 A:Accession: T45461
 A:Status: preliminary; translated from GR/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-355, 'E', 357-1005, 'H', 1007-1020, 'Q', 1021-1061, 'ERS' <HOR>
 A:Cross-references: EMBL:U48358; NID:g1203906; PIDN:AA047153.1; PID:g1203907
 A:Experimental source: tissue type ovarian
 C:Genetics:
 A:Gene: pipsqueak; psq
 A:Map position: 11
 A:Introns: 427/3
 C:Function:
 A:Description: required for establishing polarity of the developing egg chamber
 C:Superfamily: POZ domain homology
 F:21-123/Domain: POZ domain homology <POZ>

Query Match 74.7%; Score 59; DB 2; Length 1085;
 Best Local Similarity 80.0%; Pred. No. 0.37;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 HKKHGHG 11
 11111111
 DB 334 HEHNGHGHG 343

Search completed: July 6, 2001, 09:17:59
 Job time: 645 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:26:37 ; Search time 37.59 Seconds
(without alignments)
10.936 Million cell updates/sec

Title: US-09-437-912-3
Sequence: 1 GHKHHGHGCK 12

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	79	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	72	91.1	619	1 KNH2_BOVIN	P01045 bos taurus
3	72	91.1	621	1 KNH1_BOVIN	P01044 bos taurus
4	65	82.3	661	1 KNG_MOUSE	O08677 mus musculus
5	62	78.5	118	1 S109_RABIT	P50117 oryctolagus
6	62	78.5	503	1 ZNT1_MOUSE	O60738 mus musculus
7	62	78.5	507	1 ZNT1_RAT	O62720 rattus norv
8	62	78.5	693	1 CAUP_DROME	P54269 drosophila
9	61	77.2	286	1 CH36_DROME	P07182 drosophila
10	60	75.9	670	1 VGS0_HSV1	O00130 ictaluriid h
11	60	75.9	2038	1 FSH_DROME	P13709 drosophila
12	58	73.4	515	1 KE4L_CAEEL	O94797 caenorhabdi
13	57	72.2	469	1 KE4_HUMAN	O92504 homo sapien
14	56	70.9	155	1 KE4_PIG	O29175 sus scrofa
15	56	70.9	352	1 KE4_BRARE	O09pub8 brachydantio
16	56	70.9	639	1 KNG_RAT	P08934 rattus norv
17	53	67.1	196	1 SLVD_ECOLI	P30856 escherichia
18	52	65.8	476	1 KE4_MOUSE	O31125 mus musculus
19	51	64.6	1028	1 OVO_DROME	P15121 drosophila
20	51	64.6	1213	1 T2D2_DROME	O24325 drosophila
21	50	63.3	113	1 S109_RAT	P50116 rattus norv
22	50	63.3	187	1 UREE_ACPPL	O54422 actinobacill
23	50	63.3	359	1 ZNT2_RAT	O62941 rattus norv
24	50	63.3	722	1 HNM2_DROME	P22808 drosophila
25	49.5	62.7	306	1 CH38_DROME	P07183 drosophila
26	49	62.0	98	1 ZC-DICDI	P15648 dictyostell
27	49	62.0	212	1 SLVD_AERH	O07046 aeromonas h
28	49	62.0	335	1 HYPB_RHOCA	P26410 rhododactyl
29	49	62.0	449	1 CSUP_DROME	O93344 drosophila
30	49	62.0	1063	1 ELFI_DROME	P13002 drosophila
31	48	60.8	147	1 TV59_CAEEL	P50439 caenorhabdi
32	48	60.8	235	1 Y057_CAEEL	O09463 caenorhabdi
33	48	60.8	297	1 MYOD_CORTA	P21572 coturnix co

34	48	60.8	439	1	COT1_YEAST	P32798 saccharomyc
35	48	60.8	576	1	LAC1_THACU	P56193 thanatephor
36	48	60.8	633	1	SNP1_YEAST	P06782 saccharomyc
37	47	59.5	112	1	S109_MOUSE	P31725 mus musculus
38	47	59.5	374	1	GCP_STRCO	O86793 streptomyc
39	47	59.5	462	1	HEMO_HUMAN	P02790 homo sapien
40	47	59.5	606	1	GAB_DROME	P25123 drosophila
41	46	58.2	85	1	ANTF_SARPE	O08617 sarcophaga
42	46	58.2	97	1	NA2M_HUMAN	O43676 homo sapien
43	46	58.2	351	1	HRPX_PLALO	P04929 plasmodium
44	46	58.2	381	1	SELP_HUMAN	P49908 homo sapien
45	45.5	57.6	302	1	HYPB_BRAVA	Q45257 bradyrhizob

ALIGNMENTS

```

RESULT 1
ID      KNG_HUMAN
AC      P01042; P01043; STANDARD; PRT; 644 AA.
DT      21-OUL-1986 (Rel. 01, Created)
DI      01-FEB-1996 (Rel. 33, Last sequence update)
DI      01-OCT-2000 (Rel. 40, Last annotation update)
DE      KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS:
DE      BRADYKININ].
GN      KNG.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RX      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC      TISSUE=Liver;
RX      MEDLINE=85234582; PubMed=2989293;
RA      Takagaki Y., Kitamura N., Nakanishi S.;
RT      "Cloning and sequence analysis of cDNAs for human high molecular
RT      weight and low molecular weight prekininogens. Primary structures of
RT      two human prekininogens."
RT      J. Biol. Chem. 260:8601-8609(1985).
RN      [2]
RN      GENE STRUCTURE.
RX      MEDLINE=85234583; PubMed=2989294;
RA      Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RT      Nakanishi S.;
RT      "Structural organization of the human kininogen gene and a model for
RT      its evolution."
RT      J. Biol. Chem. 260:8610-8617(1985).
RN      [3]
RN      SEQUENCE OF 1-401 FROM N.A.
RX      MEDLINE=85122621; PubMed=6441591;
RA      Ohnubo T., Kurachi K., Takasawa T., Shikawa H., Sasaki M.;
RT      "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT      its identity with low molecular weight kininogen."
RT      Biochemistry 23:5691-5697(1984).
RN      [4]
RN      SEQUENCE OF 379-644.
RX      MEDLINE=86030270; PubMed=4054110;
RA      Lottspeich F., Kellermann J., Henschen A., Foerisch B.,
RT      Mueller-Esterl W.;
RT      "The amino acid sequence of the light chain of human high-molecular-
RT      mass kininogen."
RT      Eur. J. Biochem. 152:307-314(1985).
RN      [5]
RN      SEQUENCE OF 381-389.
RX      MEDLINE=90255622; PubMed=4952632;
RA      Pierce J.V.;
RT      "Structural features of plasma kinins and kininogens."
RT      Fed. Proc. 27:52-57(1968).
RN      [6]
RN      DISULFIDE BONDS.
RA      Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT      "Disulfide bonds in bovine HMW kininogens."

```

RL Selkagaku 56:808-808(1984).

CC -I- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOI. PROTEASES; (2)

CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY

CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO

CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-

CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE

CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF

CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE

CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURETIC AND

CC DIURETIC, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A

CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR

CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF

CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS

CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION),

CC INDIRECTLY VIA ENDOTHELIIUM-DERIVED RELAXING FACTOR ACTION); (5)

CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-

CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD

CC CLOTTING.

CC -I- SUBCELLULAR LOCATION: SECRETED.

CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -I- TISSUE SPECIFICITY: PLASMA.

CC -I- PFM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

CC -I- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: K02566; AAA5497.1; -

DR EMBL: M11437; AAB59550.1; -

DR EMBL: M11438; AAB59550.1; JOINED.

DR EMBL: M11521; AAB59550.1; JOINED.

DR EMBL: M11523; AAB59550.1; JOINED.

DR EMBL: M11524; AAB59550.1; JOINED.

DR EMBL: M11525; AAB59550.1; JOINED.

DR EMBL: M11526; AAB59550.1; JOINED.

DR EMBL: M11527; AAB59550.1; JOINED.

DR EMBL: M11528; AAB59550.1; JOINED.

DR EMBL: M11437; AAB59551.1; -

DR EMBL: M11438; AAB59551.1; JOINED.

DR EMBL: M11521; AAB59551.1; JOINED.

DR EMBL: M11522; AAB59551.1; JOINED.

DR EMBL: M11523; AAB59551.1; JOINED.

DR EMBL: M11524; AAB59551.1; JOINED.

DR EMBL: M11525; AAB59551.1; JOINED.

DR EMBL: M11526; AAB59551.1; JOINED.

DR EMBL: M11527; AAB59551.1; JOINED.

DR EMBL: M11528; AAB59551.1; JOINED.

DR PIR: A01279; KGHU1.

DR PIR: A25276; A25276.

DR PIR: A01280; KGHU1.

DR PIR: B25276; B25276.

DR PIR: S02482; S02482.

DR SWISS-2DPAGE; P01043; HUMAN.

DR MIM: 228960; -

DR InterPro: IPR000010; -

DR InterPro: IPR002395; -

DR Pfam: PF000031; cystatin.3.

DR PRINTS: PR00334; KININOGEN.

DR PROSITE: PS00287; CYSTATIN.2.

KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;

KW Bradykinin; Blood coagulation; Inflammatory response; Signal;

KW Alternative splicing.

FT SIGNAL 1 18

FT CHAIN 19 644 KININOGEN.

FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 389 BRADYKININ.

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD.RES	19	19	
FT	DISULFID	28	614	
FT	DISULFID	83	94	
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	205	205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	294	294	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPLIC	402	427	
FT	VARSPLIC	428	644	VSPPTSMAPADEERDSEKQGHTR -> SHLSRCEYKGR
FT	VARSPLIC	593	593	PKKGAEPASREYS (IN ISOFORM LMW).
FT	CONFLICT	593	593	MISSING (IN ISOFORM LMW).
FT	CONFLICT	593	593	T -> I (IN REF. 1).
FT	SEQUENCE	644 AA; 71945 MW; 3132B4CBAFFB7E CRC64;		

Query Match 100.0%; Score 79; DB 1; Length 644;

Best Local Similarity 100.0%; Pred. No. 9.9e-05; Mismatches 0; Gaps 0;

Matches 12; Conservative 0; Indels 0; Gaps 0;

QY 1 GHKHKHGCHGK 12

DB 494 GHKHKHGCHGK 505

RESULT 2

ID	KNH2_BOVIN	STANDARD;	PRT;	619 AA.
AC	P01045;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	KININOGEN, HMW II PRECURSOR (THIOI. PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
OX	NCBI_TaxID=9913;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84014106; Pubmed=6571699;			
RA	Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;			
RT	"A single gene for bovine high molecular weight and low molecular weight kininogens".			
RT	Weight kininogens".			
RL	Nature 305:545-549(1983).			
RN	(2)			
RP	SEQUENCE OF 19-376.			
RX	MEDLINE=87137530; Pubmed=3546295;			
RA	Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.;			

RA Miyata T., Iwanaga S.;
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion.";
 RL J. Biol. Chem. 262:2768-2779(1987).
 RN [3]
 RP SEQUENCE OF 376-391.
 RX MEDLINE=70180420; PubMed=4986212;
 RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 RT bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [4]
 RN SEQUENCE OF 387-455.
 RX MEDLINE=76260155; PubMed=956151;
 RA Han Y.N., Kato H., Iwanaga S., Suzuki T.;
 RT "Primary structure of bovine plasma high-molecular-weight kininogen.
 RT The amino acid sequence of a glycopeptide portion (fragment 1)
 RT following the C-terminus of the bradykinin moiety.";
 RL J. Biochem. 79:1201-1222(1976).
 RN [5]
 RN SEQUENCE OF 456-496.
 RX MEDLINE=75170265; PubMed=1169237;
 RA Han Y.N., Komiyama M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 RT kininogen. Amino acid sequence of a fragment ('histidine-rich
 RT peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOLE PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
 CC MATRICESIS AND DILATERS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 398.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: V01492; CAA24736.1; -;
 DR EMBL: V01492; CAA24737.1; ALT_SEQ.
 DR PIR: A01282; KGBOH2.
 DR PIR: B29559; B29559.
 DR HSSP: P04129; IAF1.
 DR InterPro: IPR000010; -;
 DR InterPro: IPR002395; -;
 DR Pfam: PF00031; cystatin; 3.
 DR PRINTS: PR00334; KININOGEN.
 DR PROSITE: PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 KW Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
 KW Inflammatory response.

FT SIGNAL 1 18
 FT CHAIN 19 619 KININOGEN, HMW II.
 FT PEPTIDE 19 376 HEAVY CHAIN.
 FT CHAIN 378 386 BRADYKININ.
 FT CHAIN 387 619 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 256 CYSTATIN-LIKE 2.
 FT DOMAIN 257 376 CYSTATIN-LIKE 3.
 FT MOD_RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 136 136 PARTIAL.
 FT CARBOHYD 168 168 OR 169.
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 400 400 INTERCHAIN.
 FT DISULFID 27 589
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 261 264
 FT DISULFID 325 337
 FT DISULFID 348 367
 FT VARIANT 398 398
 FT VARIANT 401 401
 FT VARIANT 454 454
 FT VARIANT 454 454
 SQ SEQUENCE 619 AA; 68710 MW; F04320A8BE0E0DA CRC64;
 Query Match 91.1%; Score 72; DB 1; Length 619;
 Best Local Similarity 91.7%; Pred. No. 0.001; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 1;
 QY 1 GHKHKHGHGK 12
 DB 470 GHKHKHGHGK 481
 RESULT 3
 KMH1_BOVIN
 ID KMH1_BOVIN STANDARD; PRT; 621 AA.
 AC P01044;
 DT 21-JUL-1986 (rel. 01, Created)
 DT 21-JUL-1986 (rel. 01, Last sequence update)
 DT 01-JUN-1994 (rel. 29, Last annotation update)
 DE KININOGEN, HMW I PRECURSOR (THIOLE PROTEINASE INHIBITOR) [CONTAINS:
 DE BRADYKININ].
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84014106; PubMed=6571699;
 RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
 RT "A single gene for bovine high molecular weight and low molecular
 RT weight kininogens.";
 RL Nature 305:545-549(1983).
 RN [2]
 RN SEQUENCE OF 19-378.
 RP MEDLINE=87137530; PubMed=3546295;
 RA Miyata T., Iwanaga S.,
 RA Miyata T., Iwanaga S.,
 RT "Bovine high molecular weight kininogen. The amino acid sequence,
 RT positions of carbohydrate chains and disulfide bridges in the heavy
 RT chain portion.";
 RL J. Biol. Chem. 262:2768-2779(1987).
 RN [3]
 RN SEQUENCE OF 378-393.
 RP MEDLINE=70180420; PubMed=4986212;

RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [41]
 RP SEQUENCE OF 458-498.
 RX MEDLINE:75170265; PUBMED:1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 kininogen. Amino acid sequence of a fragment ('histidine-rich
 peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975)
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C)
 CC NARURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 CC (4E) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 CC ACTION), INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 CC ACTION).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
 CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
 CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 CC TO RESIDUE 400.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONSTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: V01491; CAA24735.1; -
 CC PIR: A01281; KGB0H1.
 CC PIR: A29559; A29559.
 CC InterPro: IPR000010; -
 CC InterPro: IPR002395; -
 CC Pfam: PF00031; cystatin. 3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSTATIN. 2.
 CC Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 CC Thiol protease inhibitor; Bradykinin; Blood coagulation;
 CC Inflammatory response; Signal.
 CC
 CC SIGNAL 1 18
 CC FT CHAIN 19 621 KININOGEN, HMW I.
 CC FT CHAIN 19 378 HEAVY CHAIN.
 CC FT PEPTIDE 380 388 BRADYKININ.
 CC FT CHAIN 389 621
 CC FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 CC FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 CC FT DOMAIN 258 378 CYSTATIN-LIKE 3.
 CC FT MOD. RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 CC FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
 CC FT CARBOHYD 136 136 PARTIAL.
 CC FT CARBOHYD 168 168 OR 169.
 CC FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 CC FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 CC FT DISULFID 27 591 INTERCHAIN.
 CC FT DISULFID 82 93
 CC FT DISULFID 106 125
 CC FT DISULFID 141 144

FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;
 Query Match 91.1%; Score 72; DB 1; Length 621;
 Best Local Similarity 91.7%; Pred. No. 0.001;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GRRKHGHGSK 12
 DB 472 GRRKHGHGSK 483
 RESULT 4
 KNG_MOUSE STANDARD; PRT; 661 AA.
 AC 008677; 008676;
 DT 01-OCT-2000 (Rel. 40; Created)
 DT 01-OCT-2000 (Rel. 40; Last sequence update)
 DT 01-OCT-2000 (Rel. 40; Last annotation update)
 DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
 GN KNG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid-10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RC STRAIN=C57BL/6 x CBA; TISSUE=Liver;
 RC Takano M., Kondoh J., Yayama K., Okamoto H.;
 RT "Molecular cloning of cDNAs for mouse low- and high- molecular
 kininogen.";
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C) NATURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4E) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION),
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- SIMILARITY: CONSTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D84435; BAA19743.1; -
 CC EMBL: D84415; BAA19742.1; -
 CC MGD: MGI:1097705; Kng.
 CC InterPro: IPR000010; -

DR InterPro; IPR002395; -
 DR InterPro; IPR003243; -
 DR Pfam; PF00031; Cystatin; 3.
 DR PRINTS; PR00334; KININOGEN.
 DR PROSITE; PS00287; CYSTATIN; 1.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 661 KININOGEN.
 FT CHAIN 19 379 KININOGEN HEAVY CHAIN.
 FT CHAIN 380 388 BRADYKININ.
 FT CHAIN 389 661 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 FT DOMAIN 258 379 CYSTATIN-LIKE 3.
 FT DOMAIN 439 524 HIS-RICH.
 FT DOMAIN 28 631 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 107 125 BY SIMILARITY.
 FT DISULFID 141 144 BY SIMILARITY.
 FT DISULFID 205 217 BY SIMILARITY.
 FT DISULFID 228 247 BY SIMILARITY.
 FT DISULFID 263 266 BY SIMILARITY.
 FT DISULFID 327 339 BY SIMILARITY.
 FT DISULFID 350 369 BY SIMILARITY.
 FT CARBOHYD 82 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 401 432 VSPYIAROEERDAETEGPFGHGMLEKO -> RLIRA
 CEYKRLSKAGAPAPERQAESSQVQK (IN ISOFORM LMN).
 FT VARSPLIC 433 661 MISSING (IN ISOFORM LMN).
 FT SEQUENCE 661 AA; 73102 MW; 77460258D58796E CRC64;

Query Match
 Best Local Similarity 82.3%; Score 65; DB 1; Length 661;
 Matches 10; Conservative 83.3%; Pred. No. 0.011; 2; Indels 0; Gaps 0;

OY 1 GHKHKHGCHK 12
 DB 514 GHGSHGCHK 525

RESULT 5
 ID S109_RABIT STANDARD; PRT; 118 AA.
 AC P50117;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CALGRANULIN B (MIGRATION INHIBITORY FACTOR-RELATED PROTEIN 14)
 DE (MRP-14) (FRAGMENT).
 GN S100A9 OR MRP-14.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Neutrophils;
 RX MEDLINE=9635378; PubMed=8702688;
 RA Yang Z., de Veer M.J., Gardiner E.E., Devenish R.J., Handley C.J.,
 RA Underwood J.R., Robinson H.C.;
 RT "Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-
 RT calgranulin C when incubated with inorganic [35S]sulfate.";
 RL J. Biol. Chem. 271:19802-19809(1996).
 RN [2]
 RN SEQUENCE OF 45-82 FROM N.A.
 RC STRAIN=NEW ZEALAND WHITE;
 RX MEDLINE=94198229; PubMed=8148323;

RA Mori S., Goto K., Goto F., Mutakami K., Ohkawara S., Yoshinaga M.;
 RT "Dynamic changes in mRNA expression of neutrophils during the course
 RT of acute inflammation in rabbits.";
 RL Int. Immunol. 6:149-156(1994).
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF091849; AAC61771.1; -
 DR EMBL; D17404; BAA04227.1; -
 DR InterPro; IPR001751; -
 DR InterPro; IPR002048; -
 DR InterPro; IPR002395; -
 DR Pfam; PF01023; S-100; 1.
 DR Pfam; PF00036; eHand; 1.
 DR PRINTS; PR00334; KININOGEN.
 DR PROSITE; PS00018; EF-HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 DR Calcium-binding; Repeat.
 FT NON_TER 1 1
 FT CA_BIND 9 22 SITE I (LOW AFFINITY) (POTENTIAL).
 FT CA_BIND 53 64 SITE II (HIGH AFFINITY) (POTENTIAL).
 FT DOMAIN 103 118 2 X 8 AA TANDDEM REPEATS OF G-H-G-H-
 FT G-H-S-H.
 FT REPEAT 103 110 1.
 FT REPEAT 111 118 2.
 FT SEQUENCE 118 AA; 13292 MW; 7496118E21AD5C41 CRC64;

Query Match
 Best Local Similarity 78.5%; Score 62; DB 1; Length 118;
 Matches 9; Conservative 81.8%; Pred. No. 0.0057;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGCHK 11
 DB 105 GHGSHGCHK 115

RESULT 6
 ID ZNT1_MOUSE STANDARD; PRT; 503 AA.
 AC O60738;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ZINC TRANSPORTER 1 (ZNT-1).
 DE SLC30A1 OR ZNT1.
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9518868; PubMed=7882967;
 RA Palmiter R.D., Findley S.D.;
 RT "Cloning and functional characterization of a mammalian zinc
 RT transporter that confers resistance to zinc.";
 RL EMBO J. 14:639-649(1995).
 CC -1- FUNCTION: MAY BE INVOLVED IN ZINC TRANSPORT OUT OF THE CELL.
 CC LETHALITY OF KNOCKOUT EARLY IN GESTATION SUGGESTS A ROLE OF THE
 CC PROTEIN IN FETAL ZINC ACQUISITION AND RETENTION.
 CC -1- SUBUNIT: MULTIMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC LOCALIZED ON THE PLASMA MEMBRANE (PROBABLE).
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.

CC -1- SIMILARITY: BELONGS TO THE SLC30A FAMILY OF TRANSPORTERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: U17132; AAA79233.1; -
CC MGD: MGI:1345281; SLC30A1.
CC InterPro: IPR002524; -
CC Pfam: PF01545; Cation_efflux; 1.
CC zinc; Transport; Transmembrane; Multigene family; Repeat.
CC
CC FT TRANSMEM 1 10 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 11 31 POTENTIAL.
CC FT TRANSMEM 32 35 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 36 56 POTENTIAL.
CC FT TRANSMEM 57 78 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 79 99 POTENTIAL.
CC FT TRANSMEM 100 113 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 114 134 POTENTIAL.
CC FT TRANSMEM 135 243 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 244 264 POTENTIAL.
CC FT TRANSMEM 265 303 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 304 324 POTENTIAL.
CC FT TRANSMEM 325 503 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 145 156 6 X 2 AA APPROXIMATE REPEATS OF H-G.
CC FT CAROHD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT SEQUENCE 503 AA; 54716 MW; 7CAFF93FC13CDA22 CRC64;

Query Match 78.5%; Score 62; DB 1; Length 503;
Best Local Similarity 81.8%; Pred. No. 0.024; 2; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGHC 11
DB 145 GHGSHGHGHC 155

RESULT 7
ZNT1_RAT STANDARD; PRT; 507 AA.
AC Q62720;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 40, Last annotation update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ZINC TRANSPORTER 1 (ZNT-1).
GN SLC30A1 OR ZNT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95188868; PubMed=7882967;
RA Palmer R.D., Findley S.D.;
RT "Cloning and functional characterization of a mammalian zinc
RT transporter that confers resistance to zinc.";
RL EMBL J. 14:639-649(1995).
RN [2]
RP INDUCTION BY ZINC.
RC TISSUE=Intestine;
RX MEDLINE=96226729; PubMed=9560190;
RA McMahon R.J., Cousins R.J.;
RT "Regulation of the zinc transporter ZNT-1 by dietary zinc.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4841-4846(1998).
CC -1- FUNCTION: MAY BE INVOLVED IN ZINC TRANSPORT OUT OF THE CELL.
CC -1- SUBUNIT: MULTIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC LOCALIZED ON THE PLASMA MEMBRANE (PROBABLE). LOCALIZED ON THE
CC BASOLATERAL SURFACE OF THE ENTEROCYTES.
CC
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE PROTEIN IS DETECTED IN
CC DUODENUM AND JEJUNUM BUT NOT IN ILEUM AND COLON.
CC
CC -1- INDUCTION: SLIGHTLY BY ZINC IN THE INTESTINE, BUT NOT THE LIVER.
CC
CC -1- SIMILARITY: BELONGS TO THE SLC30A FAMILY OF TRANSPORTERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: U17133; AAA79234.1; -
CC InterPro: IPR002524; -
CC Pfam: PF01545; Cation_efflux; 1.
CC zinc; Transport; Transmembrane; Multigene family; Repeat.
CC
CC FT TRANSMEM 1 10 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 11 31 POTENTIAL.
CC FT TRANSMEM 32 35 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 36 56 POTENTIAL.
CC FT TRANSMEM 57 78 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 79 99 POTENTIAL.
CC FT TRANSMEM 100 113 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 114 134 POTENTIAL.
CC FT TRANSMEM 135 247 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 248 268 POTENTIAL.
CC FT TRANSMEM 269 307 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 308 328 POTENTIAL.
CC FT TRANSMEM 329 507 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 145 156 6 X 2 AA APPROXIMATE REPEATS OF H-G.
CC FT CAROHD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT SEQUENCE 507 AA; 55142 MW; 9F9770017C2455FC CRC64;

Query Match 78.5%; Score 62; DB 1; Length 507;
Best Local Similarity 81.8%; Pred. No. 0.024; 2; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGHC 11
DB 145 GHGSHGHGHC 155

RESULT 8
CAUP_DROME STANDARD; PRT; 693 AA.
AC P54269;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HOMEOBOX PROTEIN CAUPOLOCAN.
GN CAUP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96180722; PubMed=8620542;
RA Gomez-Skarmeta J.-L., del Corral R.D., de la Calle-Mustienes E.,
RA Ferrer-Marco D., Modolell J.;
RT "Araucan and caupolican, two members of the novel iroquois complex,
RT encode homeoproteins that control proneural and vein-forming genes.";
RL Cell 85:95-110(1996).
CC -1- FUNCTION: CONTROLS PRONEURAL AND VEIN FORMING GENES. POSITIVE
CC TRANSCRIPTIONAL CONTROLLER OF AC-SC (ACHAETE-SCUTE). MAY ACT AS AN
CC ACTIVATOR THAT INTERACTS WITH THE TRANSCRIPTIONAL COMPLEX
CC ASSEMBLED ON THE AC AND SC PROMOTERS AND PARTICIPATES IN
CC TRANSCRIPTION INITIATION.

CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE TALE/TRO FAMILY OF HOMEOBOX PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; X95178; CAAG4485.1; -.
CC HSSP; P02833; ISAN.
CC FlyBase; FBgn0015919; caup.
CC InterPro; IPR001356; -.
CC Pfam; PF000046; homeobox; 1.
CC PROSITE; PS00027; HOMEOBOX_1; 1.
CC PROSITE; PS50071; HOMEOBOX_2; 1.
CC Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
CC Developmental protein.
CC DNA_BIND 226 288 HOMEOBOX (TALE-TYPE).
FT DOMAIN 300 303 POLY-ASP.
FT DOMAIN 405 418 POLY-GLN.
FT DOMAIN 501 516 POLY-GLN.
FT DOMAIN 517 528 POLY-HIS.
FT DOMAIN 565 572 POLY-SER.
FT DOMAIN 613 624 POLY-SER.
SQ SEQUENCE 693 AA; 73749 MW; 8E0D6D43C9CDC619 CRC64;

Query Match 78.5%; Score 62; DB 1; Length 693;
Best Local Similarity 81.8%; Pred. No. 0.032;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GHKHKHGHC 11
Db 649 GHGSHGHG 659

RESULT 9
CH36_DROME STANDARD; PRT; 286 AA.
ID CH36_DROME STANDARD; PRT; 286 AA.
AC P07182;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE CHORION PROTEIN S36.
GN CP36 OR S36.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CC Ephydroidea; Drosophilidae; Drosophila.
CC NCBI_TaxID=7227;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=87246506; Pubmed=3036489;
CC Spreading A.C., de Cicco D.V., Makimoto B.T., Levine J.F.,
CC Kalfayan L.J., Cooley L.;
CC "Amplification of the X-linked Drosophila chorion gene cluster
CC requires a region upstream from the s38 chorion gene.";
CC EMBO J. 6:1045-1053(1987).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; X05245; CAA28670.1; -.
CC PIR; S07193; S07193.
CC FlyBase; FBgn0000359; CP36.
CC SEQUENCE 670 AA; 64174 MW; 2B64A781C519E8B4 CRC64;

KW Chorion.
SQ SEQUENCE 286 AA; 30080 MW; 29B7CC12E1F1833 CRC64;

Query Match 77.2%; Score 61; DB 1; Length 286;
Best Local Similarity 75.0%; Pred. No. 0.019;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GHKHKHGHC 12
Db 27 GHGSHGHG 38

RESULT 10
VG50_HSV11 STANDARD; PRT; 670 AA.
ID VG50_HSV11 STANDARD; PRT; 670 AA.
AC Q00130;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL GENE 50 PROTEIN.
GN 50.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC unclassified Herpesviridae.
CC NCBI_TaxID=10401;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=AUBURN 1;
CC MEDLINE=92087490; Pubmed=1727613;
CC Davidson A.J.;
CC "Channel catfish virus: a new type of herpesvirus.";
CC Virology 186:9-14(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; M75136; AAA88153.1; -.
CC PIR; F36791; F36791.
CC Hypothetical protein; Repeat.
CC REPEAT 143 158
FT REPEAT 171 186
FT REPEAT 200 214
FT REPEAT 215 233
FT REPEAT 234 252
FT REPEAT 253 268
FT REPEAT 279 293
FT REPEAT 294 309
FT REPEAT 320 334
FT REPEAT 335 349
FT REPEAT 362 376
FT REPEAT 377 391
FT REPEAT 392 406
FT REPEAT 407 421
FT REPEAT 422 436
FT REPEAT 437 452
FT REPEAT 464 477
FT REPEAT 478 493
FT REPEAT 504 517
FT REPEAT 518 531
FT REPEAT 532 545
FT REPEAT 546 559
FT REPEAT 560 573
FT REPEAT 574 587
FT REPEAT 588 601
FT REPEAT 602 615
FT REPEAT 616 629
SQ SEQUENCE 670 AA; 64174 MW; 2B64A781C519E8B4 CRC64;

Query Match 75.9%; Score 60; DB 1; Length 670;
 Best Local Similarity 81.8%; Pred. No. 0.062;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHHGHGHC 11
 ||| |||||
 DB 636 GHGHGHGHC 646

RESULT 11

FSH_DROME STANDARD: PRT: 2038 AA.
 ID FSH_DROME P13709; P13710;
 AC 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE FEMALE STERILE HOMEOCTIC PROTEIN (FRAGILE-CHORION MEMBRANE PROTEIN).
 GN FS(1)H OR FSH.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89276730; PubMed=2567251;
 RA Haynes S.R., Mozer B.A., Bhatia-Dey N., David I.B.;
 RT "The Drosophila fsh locus, a maternal effect homeotic gene, encodes
 RT apparent membrane proteins.";
 RL Dev. Biol. 134:246-257(1989).
 CC -1- FUNCTION: REQUIRED MATERNALLY FOR PROPER EXPRESSION OF OTHER
 CC HOMEOCTIC GENES INVOLVED IN PATTERN FORMATION, SUCH AS UBX.
 CC -1- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.
 CC -1- SIMILARITY: CONTAINS 2 BROMODOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 ET DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M23221; AAA28540.1; -;
 DR EMBL: M23222; AAA28541.1; ALT_TERM.
 DR EMBL: M15762; AAA70424.1; -;
 DR EMBL: M15763; AAA70423.1; -;
 DR EMBL: M15764; AAA70422.1; -;
 DR PIR: A43742; A43742.
 DR HSSP: P04002; 1WFA.
 DR FLYBase: FBgn0004656; fs(1)h.
 DR InterPro: IPR001487; -;
 DR Pfam: PF00439; Bromodomain; 2.
 DR PRINTS: PR00503; Bromodomain.
 DR PROSITE: PS00633; BROMODOMAIN_1; 2.
 DR PROSITE: PS50014; BROMODOMAIN_2; 2.
 DR Developmental protein; Bromodomain; Transmembrane; Repeat.
 KW DOMAIN 51 123 BROMODOMAIN 1.
 FT DOMAIN 495 567 BROMODOMAIN 2.
 FT DOMAIN 945 1106 ET DOMAIN.
 FT TRANSMEM 330 350 POTENTIAL.
 FT TRANSMEM 451 471 POTENTIAL.
 FT TRANSMEM 750 770 POTENTIAL.
 FT TRANSMEM 790 810 POTENTIAL.
 FT TRANSMEM 816 830 POTENTIAL.
 FT TRANSMEM 874 894 POTENTIAL.
 FT TRANSMEM 1731 1751 POTENTIAL.
 FT TRANSMEM 1939 1959 POTENTIAL.
 FT VARIANT 909 909 G -> A.
 FT VARIANT 1022 1022 H -> RRPY.

SQ SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match 75.9%; Score 60; DB 1; Length 2038;
 Best Local Similarity 81.8%; Pred. No. 0.18;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHHGHGHC 11
 ||| |||||
 DB 597 GHGHGHGHC 607

RESULT 12

KE4L_CAEEL STANDARD: PRT: 515 AA.
 ID KE4L_CAEEL Q9XT07;
 AC 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL KE4-LIKE PROTEIN H13N06.5 IN CHROMOSOME X.
 GN H13N06.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Lennard N.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE KE4/CATSP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z99942; CAB17070.1; -;
 DR WormPep: H13N06.5; CE18815.
 KW Hypothetical protein; Transmembrane; Glycoprotein.
 FT TRANSMEM 49 47 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT TRANSMEM 247 267 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 386 406 POTENTIAL.
 FT TRANSMEM 429 449 POTENTIAL.
 FT TRANSMEM 463 483 POTENTIAL.
 FT TRANSMEM 92 182 POTENTIAL.
 FT DOMAIN 7 7 HIS-RICH.
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 515 AA; 55500 MW; 17D7E854F4E1DAAF CMC64;
 Query Match 73.4%; Score 58; DB 1; Length 515;
 Best Local Similarity 72.7%; Pred. No. 0.093;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

RESULT 13
 KE4_HUMAN STANDARD: PRT: 469 AA.
 ID KE4_HUMAN Q92504; Q9UI00;
 AC 276 GHSHGHSHGSHG 286

01-OCT-2000 (Rel. 40, Created)
 01-OCT-2000 (Rel. 40, Last sequence update)
 01-OCT-2000 (Rel. 40, Last annotation update)
 HISTIDINE-RICH MEMBRANE PROTEIN KE4.
 HKE4 OR RINGS.
 Homo sapiens (Human).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 [1]
 TISSUE: FROM N.A.
 MEDLINE=97001166; PubMed=8612499;
 Ando A., Kikuta Y., Shigenari A., Kawata H., Okamoto N., Shina T.,
 Chen L., Ikemura T., Abe K., Kimura M., Inoko H.;
 "cDNA cloning of the human homologues of the mouse Ke4 and Ke6 genes
 at the centromeric end of the human MHC region.";
 Genomics 35:600-602(1996).
 [2]
 SEQUENCE FROM N.A.
 Vergara A., Lana I., Corella A., de Miguel C., Miglaccio M.,
 Encio I.;
 "Molecular cloning and characterization of the human KE4 gene and 5'
 flanking region.";
 Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 SEQUENCE FROM N.A.
 Tubby B.;
 Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 -1- TISSUE SPECIFICITY: MAJOR EXPRESSION IN PLACENTA, LUNG, KIDNEY
 AND PANCREAS.
 -1- SIMILARITY: BELONGS TO THE KE4/CATSP FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

 CC EMBL: D82060; BAA1528.1; -
 DR EMBL: AF117221; AAD12305.1; -
 DR EMBL: AL031228; CAA20238.1; -
 KW MIM: 601416; -
 Transmembrane; Glycoprotein.
 FT TRANSMEM 10 30
 FT TRANSMEM 138 158
 FT TRANSMEM 169 189
 FT TRANSMEM 214 234
 FT TRANSMEM 381 401
 FT TRANSMEM 417 436
 FT DOMAIN 30 114
 FT DOMAIN 238 263
 FT CAROHD 330 330
 FT CONFLICT 7 7
 FT CONFLICT 280 280
 FT CONFLICT 376 376
 FT CONFLICT 397 469
 SEQUENCE 469 AA: 50118 MW: 6504A1EF5A6A5B9 CXC64;
 AND 2).
 Query Match 72.2%; Score 57; DB 1; Length 469;
 Best Local Similarity 80.0%; Pred. No. 0.12;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 GHKHHGHGH 10
 ||| |||||
 DB 57 GHSHGHGH 66

RESULT 14
 ID KE4_PIG STANDARD; PRT: 155 AA.
 AC Q29175; G9XT01;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HISTIDINE-RICH PROTEIN KE4 (FRAGMENTS).
 HKE4.
 Sus scrofa (Pig).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 NCBI_TaxID=96823;
 [1]
 SEQUENCE OF 1-124 FROM N.A.
 RC TISSUE=Small intestine;
 RX MEDLINE=96327607; PubMed=8672129;
 Wintrobe A.K., Fredholm M., Davies W.;
 "Evaluation and characterization of a porcine small intestine cDNA
 library: analysis of 839 clones.";
 Mamm. Genome 7:509-517(1996).
 [2]
 SEQUENCE OF 125-155 FROM N.A.
 RC STRAIN=Belgian landrace;
 RA Chardon P., Rogel-Galliard C., Peelman L.J., Yerle M., Renard C.,
 Vaiman M.;
 "Physical organization of the swine major histocompatibility complex
 class II region.";
 Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 -1- SIMILARITY: BELONGS TO THE KE4/CATSP FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

 CC EMBL: F14787; CAA23256.1; -
 DR EMBL: AF146397; AAD44801.1; -
 KW Transmembrane.
 FT NON_TER 1 1
 FT TRANSMEM 19 39
 FT DOMAIN 49 59
 FT DOMAIN 81 86
 FT NON_CONS 124 125
 FT NON_TER 155 155
 SEQUENCE 155 AA: 16840 MW: 56354DE3CAAE4524 CXC64;
 Query Match 70.9%; Score 56; DB 1; Length 155;
 Best Local Similarity 80.0%; Pred. No. 0.057;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 GHKHHGHGH 10
 ||| |||||
 DB 48 GHSHGHGH 57
 RESULT 15
 ID KE4_BRAE STANDARD; PRT: 352 AA.
 AC Q9PUB8;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HISTIDINE-RICH MEMBRANE PROTEIN KE4 HOMOLOG (FRAGMENT).
 HKE4.
 Brachydanio rerio (Zebrafish) (Zebrafish).
 OS

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.
 OX NCBI_TaxID=7955;

RN [1]
 RP SEQUENCE FROM N.A.

RA Murray B.W., Suelmann H., Klein J.,
 RT "Identification of a homolog of the human HKE4 gene in zebrafish."
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE KEA/CATSUP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: AF196345; AAF05821.1; -
 KM Transmembrane: Glycoprotein.

FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 128 148 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT DOMAIN 24 105 HIS-RICH.
 FT DOMAIN 177 217 HIS-RICH.
 FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 352 352
 SQ SEQUENCE 352 AA: 37922 MW: C8C8C60F6D2BA8A6 CRC64;

Query Match 70.9%; Score 56; DB 1; Length 352;

Best Local Similarity 80.0%; Pred. No. 0.13;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GAKKHHGH 10
 |||||||

DB 94 GAKKHHGH 103

Search completed: July 6, 2001, 09:26:37
 Job time: 968 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:25:52 ; Search time 118.42 Seconds
(without alignments)
13.407 Million cell updates/sec

Title: US-09-437-912-3

Perfect score: 79

Sequence: 1 GHKHHGHGHC 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.todent:*
- 12: sp.unclassified:*
- 13: sp.vertebrate:*
- 14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	84.8	450	5	027920 bradydia hy
2	64	81.0	86	10	022671 alnus glut
3	64	81.0	99	10	092RC7
4	62	78.5	189	2	09PU12
5	62	78.5	398	10	081036
6	62	78.5	490	2	09X9W6
7	62	78.5	494	5	09GRW9
8	62	78.5	495	5	09VEX1
9	62	78.5	554	5	09W4C1
10	62	78.5	693	5	09VU00
11	61	77.2	79	10	09M435
12	61	77.2	213	5	09GTNO
13	61	77.2	457	5	09W416
14	61	77.2	605	5	077280
15	60	75.9	206	5	09GTN1
16	60	75.9	218	5	09V3P9
17	60	75.9	385	5	09VWX5
18	60	75.9	686	5	09VWS0
19	60	75.9	1937	5	09W313

20	59	74.7	1064	5	09V5N1	09v5n1 drosophila
21	59	74.7	1085	5	024455	024455 drosophila
22	58	73.4	126	11	009016	009016 rattus norv
23	58	73.4	519	5	09VU19	09vul9 drosophila
24	57	72.2	318	3	09P5K2	09p5k2 neurospora
25	57	72.2	325	5	09VYM9	09vym9 drosophila
26	56	70.9	348	5	09W2X1	09w2x1 drosophila
27	56	70.9	1150	5	09VWN4	09vwn4 drosophila
28	56	70.9	1235	5	09VEF3	09vef3 drosophila
29	56	70.9	1493	5	09VEF7	09vef7 drosophila
30	55	69.6	198	5	09NNV9	09nnv9 plasmodium
31	55	69.6	224	2	069955	069955 streptomyce
32	55	69.6	336	2	055451	055451 synechocyst
33	55	69.6	378	10	09LX81	09lxs1 arabidopsis
34	55	69.6	386	10	09FPW7	09fpw7 thlaspi cae
35	55	69.6	492	11	09UKN2	09ukn2 mus musculu
36	55	69.6	989	5	09W2S4	09w2s4 drosophila
37	54	68.4	199	10	09LYB2	09lyb2 arabidopsis
38	54	68.4	251	5	018577	018577 caenorhabdi
39	54	68.4	334	10	09M271	09m271 arabidopsis
40	54	68.4	337	5	09VHP7	09vhp7 drosophila
41	54	68.4	344	10	082643	082643 arabidopsis
42	54	68.4	612	2	09K2U9	09k2u9 streptomyce
43	54	68.4	1140	5	09VD60	09vd60 drosophila
44	54	68.4	1221	5	024079	024079 drosophila
45	54	68.4	1891	5	077275	077275 drosophila

ALIGNMENTS

RESULT 1						
027920		PRELIMINARY:	PRT:	450 AA.		
AC 027920:						
DT 01-NOV-1996 (TREMBLrel. 01, Created)						
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)						
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)						
DE PC4-1.						
GN PC4-1.						
OS Bradydia hygida.						
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;						
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Scleroidea;						
OC Scleridae; Bradydia.						
OX NCBI_TaxID=35572;						
RN [1]						
RP SEQUENCE FROM N.A.						
RC TISSUE-SALIVARY GLAND;						
RX MEDLINE=95393845; PubMed=7664619;						
RA Monesi N., Fernandez M.A., Fontes A.M., Basso L.R., Nakanishi Y.,						
RA Baron B., Butlin G., Peco-Larson M.L.;						
RT Molecular characterization of an 18 kb segment of DNA puf C4 of						
RT Bradydia hygida (Diptera, scleridae).";						
RL Chromosoma 103:715-724(1995).						
DR EMBL: U13883; AAA83554.1; -;						
DR EMBL: U13892; AAA83555.1; -;						
SQ SEQUENCE 450 AA; 47185 MW; 1F0633CE9B7F964C CRC64;						
Query Match	84.8%;	Score 67;	DB 5;	Length 450;		
Best Local Similarity	90.9%;	Pred. No. 0.0087;				
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;						
OY 1 GHKHHGHGHC 11						
DB 76 GHKHHGHGHC 86						
RESULT 2						
022671		PRELIMINARY:	PRT:	86 AA.		
AC 022671:						
DT 01-JAN-1998 (TREMBLrel. 05, Created)						

DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
 DE AG164 PROTEIN PRECURSOR.
 GN AG164.
 OS Alnus glutinosa (Alder).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97348585; PubMed=9204569;
 RA Pawlowski K., Twigg P., Dobritsa S., Guan C., Mullin B.C.;
 RT "A nodule-specific gene family from Alnus glutinosa encodes glycine-
 and histidine-rich proteins expressed in the early stages of
 actinorhizal nodule development."
 RL MOL. Plant Microbe Interact. 10:656-664(1997).
 DR EMBL: Y08436; CAA69708.1; -.
 KW Signal.
 FT CHAIN 30 86 POTENTIAL.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 86 POTENTIAL.
 SQ SEQUENCE 86 AA; 9188 MW; D85B7EF888C8899A CRC64;

Query Match 81.0%; Score 64; DB 10; Length 86;
 Best Local Similarity 81.8%; Pred. No. 0.0048;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GHKHHGHGHC 11
 DB 51 GHRHVHGHGHC 61

RESULT 3
 ID 092RC7 PRELIMINARY; PRT; 99 AA.
 AC 092RC7;
 DT 01-MAY-1999 (TRENBLREL. 10, Created)
 DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
 DE ACTINORHIZAL NODULIN AGNOD-GHRP.
 GN AGN84.
 OS Alnus glutinosa (Alder).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Dobritsa S.V., Mullin B.C.;
 RT "In vitro expression of actinorhizal nodulin AGNOD-GHRP and
 demonstration of its toxicity to Escherichia coli."
 RL (in) Stacey G., Mullin B.C., Gresshoff P.M. (eds.);
 RL the Biology of Plant-Microbe Interactions:
 RL Proceedings of the 8th International Symposium on Molecular
 RL Plant-Microbe Interactions, pp.1-1, Unknown Publisher (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Twigg P.G.;
 RT "Isolation of a nodule-specific cDNA encoding a putative glycine-rich
 protein from Alnus glutinosa."
 RL Thesis (1993), The University of Tennessee, Knoxville, TN, USA.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Pawlowski K., Twigg P.G., Dobritsa S.V., Guan C., Mullin B.C.;
 RT "A nodule-specific gene family from Alnus glutinosa encodes glycine
 and histidine-rich proteins expressed in the early stages of
 actinorhizal nodule development."
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U69156; AAD00171.1; -.

DR InterPro: IPR002395; -.
 DR PRINTS: PR00334; KININOGEN.
 SQ SEQUENCE 99 AA; 10567 MW; 2ACBE4D57C070E83 CRC64;

Query Match 81.0%; Score 64; DB 10; Length 99;
 Best Local Similarity 81.8%; Pred. No. 0.0056;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GHKHHGHGHC 11
 DB 51 GHRHVHGHGHC 61

RESULT 4
 ID 09PJ12 PRELIMINARY; PRT; 189 AA.
 AC 09PJ12;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
 DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE.
 GN SLXD OR C30115.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCT 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vleet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences."
 RL Nature 403:665-668(2000).
 RL Nature 403:665-668(2000).
 DR EMBL: AL139074; CAB72599.1; -.
 DR InterPro: IPR001179; -.
 DR PROSITE: P550059; FKBP_PPIASE_3; 2.
 SQ SEQUENCE 189 AA; 20132 MW; 47B5F5D047549D7F CRC64;

Query Match 78.5%; Score 62; DB 2; Length 189;
 Best Local Similarity 81.8%; Pred. No. 0.021;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHHGHGHC 11
 DB 166 GHDHGHGHC 176

RESULT 5
 ID 081036 PRELIMINARY; PRT; 398 AA.
 AC 081036;
 DT 01-NOV-1998 (TRENBLREL. 08, Created)
 DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
 DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
 DE PUTATIVE ZINC TRANSPORTER.
 GN F19D11.8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;

RT "Arabidopsis thaliana chromosome II BAC F19D11 genomic sequence."
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC005310; AAC33498.1; -
 DR InterPro: IPR002524; -
 DR Pfam: PF01545; Cation_efflux; 1.
 SQ SEQUENCE 398 AA; 43827 MW; 7E20E0B29237BB23 CRC64;

Query Match
 Best Local Similarity 78.5%; Score 62; DB 10; Length 398;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGCHG 11
 ||| |
 DB 185 GHSHGHGCHG 195

RESULT 6
 O9X9M6 PRELIMINARY; PRT; 490 AA.
 AC O9X9M6:
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE HYPOTHETICAL 52.7 KDA PROTEIN.
 GN SC17.24C.
 OS Streptomyces coelicolor
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 NC NCB1_TaxID=1902;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K., Harris D.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleser H.M., Denapalte D., Elchner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mcl. Microbiol. 21:77-96(1996).
 DR EMBL: AL096743; CAB46407.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 490 AA; 52666 MW; 6CCBAD04D68E50A CRC64;

Query Match
 Best Local Similarity 78.5%; Score 62; DB 2; Length 490;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGCHG 11
 ||| |
 DB 49 GHSHGHGCHG 59

RESULT 7
 O9GRW9 PRELIMINARY; PRT; 494 AA.
 AC O9GRW9:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE ZINC/IRON REGULATED TRANSPORTER-RELATED PROTEIN 3, ZIP3 PROTEIN.

GN ZIP3.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON S.
 RA Lind M.I., Soderhall K.;
 RT "Identification of three zinc-iron regulated transporter-like genes in
 Drosophila melanogaster."
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ401615; CAC14874.1; -
 SQ SEQUENCE 494 AA; 54201 MW; 2FB595C6E3BE8B8 CRC64;

Query Match
 Best Local Similarity 78.5%; Score 62; DB 5; Length 494;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKHKHGCHG 11
 ||| |
 DB 236 GHSHGHGCHG 246

RESULT 8
 O9VEX1 PRELIMINARY; PRT; 495 AA.
 AC O9VEX1:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CG6898 PROTEIN.
 GN ZIP3 OR CG6898.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelink S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 ABail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brostein P., Brotlier P.,
 RA Burlis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Galbert W.M., Glaser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaitali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Splier E, Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zdzveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
DR EMBL: AE003712; AAF55295.1.
DR FlyBase: FBgn0036412; CG6898.
SQ SEQUENCE 495 AA: 54292 MW: AAIDBEE85446FFCB CRC64:

Query Match	78.5%	Score 62	DB 5	Length 495
Best Local Similarity	81.8%	Pred. No. 0.055		
Matches	9	Conservative	0	Mismatches 2
				Indels 0
				Gaps 0
QY	1	GHKRKHGHG 11		
db	237	GCHGSHGCHG 247		

RESULT	9			
ID	Q9WAC1	PRELIMINARY:	PRT:	554 AA.
AC	Q9WAC1;			
DT	01-MAY-2000 (TREMBlrel. 13, Created)			
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)			
DT	01-JUN-2000 (TREMBlrel. 14, Last annotation update)			
DE	CG15784 PROTEIN.			
GN	CG15784			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pretler B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,			
RA	Abdl J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,			
RA	Ballew R.M., Basu A.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeoon K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Hock J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,			
RA	Jakali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P.E., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy J., Munzy D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,			
RA	Palazzo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Stavitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,			

	Query Match	Score 62	DB 5	Length 554
	Best Local Similarity	75.0%	Pred. No. 0.061	
	Matches 9	Conservative 1	Mismatches 2	Indels 0
				Gaps 0
Oy	1 GHRKRGHGCK 12			
db	542 GGGHGCHGHR 553			

RESULT	ID	PRELIMINARY:	PRT:	693 AA.
09VU00	09VU00			
AC	09VU00;			
DT	01-MAY-2000 (TREMBLREL. 13, Created)			
DT	01-MAY-2000 (TREMBLREL. 13, Last sequence update)			
DT	01-MAR-2001 (TREMBLREL. 16, Last annotation update)			
DE	CAUP PROTEIN			
GN	CAUP OR CG10605.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxId=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BERKELEY;			
RX	MEDLINE=20196006; PubMed=107311132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,			
RA	Abbil J.F., Abghyani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,			
RA	Balcer R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Besson K.Y., Benos P.V., Betman B.P., Bhandal D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,			
RA	Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Liao X.P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,			
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Soler E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Styrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissendbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER HOMEOBOX DOMAINS.
 DR EMBL: AE003540; AAF49895.1; -
 DR Flybase: FBgn0015919; caup.
 DR InterPro: IPR001356; -
 DR Pfam: PF00046; homeobox_1.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS50071; HOMEOBOX_2; 1.
 DR SMART: SM00389; HOX; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 693 AA; 73667 MW; FBBE1616493F7EC9 CRC64;

Query Match 78.5%; Score 62; DB 5; Length 693;
 Best Local Similarity 81.8%; Pred. No. 0.076;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GHKHKHGCHG 11
 1111111111
 DB 649 GHGSHGHGCHG 659

RESULT 11
 Q9M435 PRELIMINARY; PRT; 79 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PHASE-CHANGE RELATED PROTEIN PRECURSOR.
 OS Eukaryota; (English oak).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1;
 OC Fagales; Fagaceae; Quercus.
 OX NCBI_TaxID=38942;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-IN VITRO SHOOT CULTURES;
 RA Gil B., Pastoriza E.M., Ballester A., Sanchez C.;
 RT "Identification of a phase-change related mRNA in oak shoot cultures
 RT derived from basal sprouts and crown branches.";
 RT Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ271778; CAB/2442.1; -
 DR Signal.
 KW Signal.
 FT SIGNAL 1 29 POTENTIAL.
 SQ SEQUENCE 79 AA; 8414 MW; 8E45CABF40F00B6F CRC64;

Query Match 77.2%; Score 61; DB 10; Length 79;
 Best Local Similarity 81.8%; Pred. No. 0.013;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GHKHKHGCHG 11
 1111111111
 DB 45 GHGSHGHGCHG 55

RESULT 12
 Q9GTNO PRELIMINARY; PRT; 213 AA.
 AC Q9GTNO;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE D506238.4-LIKE PROTEIN (FRAGMENT).
 OS *Drosophila yakuba* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schmid K.J., Aquadro C.F.;
 RT "The evolutionary analysis of 'orphans' from the *Drosophila* genome
 RT identifies incorrectly annotated and rapidly evolving genes.";
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF264920; MAG10259.1; -
 DR NON_TER 213 213
 FT SEQUENCE 213 AA; 23731 MW; 315B8590D978B6C9 CRC64;

Query Match 77.2%; Score 61; DB 5; Length 213;
 Best Local Similarity 81.8%; Pred. No. 0.034;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GHKHKHGCHG 11
 1111111111
 DB 180 GHGSHGHGCHG 190

RESULT 13
 Q9M416 PRELIMINARY; PRT; 457 AA.
 AC Q9M416;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE EG:84H4.4 OR CG3081.
 GN EG:84H4.4 OR CG3081.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaitani B.E., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodege T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003431; AAF45965.1; -;
 FT FLYBASE: FBgn0025613; EG:84H4.4.
 SQ SEQUENCE 437 AA; 48919 MW; 70B4B7ADDD02E0AD CRC64;

Query Match 77.2%; Score 61; DB 5; Length 457;
 Best Local Similarity 81.8%; Pred. No. 0.071; 2; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKKHGHGHG 11
 ||| |||||
 DB 340 GHCHHGHGHG 350

RESULT 14

ID 077280 PRELIMINARY; PRT; 605 AA.
 AC 077280;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE EG:84H4.4 PROTEIN.
 GN EG:84H4.4 OR CG3081.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ferraz C., Vidal S., Brun C., Bucheton A., Demallie J.G.;
 RT "Sequencing the distal X chromosome of *Drosophila melanogaster*.";
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Benos P.;
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL031766; CAA21135.1; -;
 DR FLYBASE: FBgn0025613; EG:84H4.4.
 SQ SEQUENCE 605 AA; 64947 MW; B06C84ACAD7D2C84 CRC64;

Query Match 77.2%; Score 61; DB 5; Length 605;
 Best Local Similarity 81.8%; Pred. No. 0.094; 2; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKKHGHGHG 11
 ||| |||||
 DB 340 GHCHHGHGHG 350

RESULT 15

ID 09GTN1 PRELIMINARY; PRT; 206 AA.
 AC 09GTN1;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE DS06238.4-LIKE PROTEIN (FRAGMENT).
 OS *Drosophila simulans* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NORTH CAROLINA 17;
 RA Schmidt K.J., Aquadro C.F.;

RT "The evolutionary analysis of 'orphans' from the *Drosophila* genome
 RT identifies incorrectly annotated and rapidly evolving genes.";
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF264919; AAC10258.1; -;
 FT NON_TER 206
 SQ SEQUENCE 206 AA; 23011 MW; E24B8C189BC0746D CRC64;

Query Match 75.9%; Score 60; DB 5; Length 206;
 Best Local Similarity 81.8%; Pred. No. 0.046; 2; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GHKKHGHGHG 11
 ||| |||||
 DB 173 GHVGHGHGHG 183

Search completed: July 6, 2001, 09:25:53
 Job time: 989 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:17 ; Search time 113.68 Seconds
(without alignments)
6.399 Million cell updates/sec

Title: US-09-437-912-4

Perfect score: 74

Sequence: 1 GKKNKGNKMK 12

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	100.0	12	21	AAV81995
2	74	100.0	16	21	AAV806337
3	74	100.0	28	21	AAV81997
4	74	100.0	94	21	AAV93351
5	74	100.0	179	21	AAV93353
6	74	100.0	186	21	AAV93349
7	74	100.0	255	21	AAV93342
8	69	93.2	47	21	AAV93345
9	69	93.2	62	21	AAV93348
10	69	93.2	63	16	AAV75186
11	69	93.2	83	21	AAV93347

12	69	93.2	131	16	AAV75181
13	46	62.2	41	16	AAV75180
14	46	62.2	110	16	AAV75178
15	41	55.4	19	21	AAV71879
16	41	55.4	20	17	AAV07625
17	40	54.1	376	18	AAV35386
18	40	54.1	376	18	AAV35387
19	39	52.7	60	16	AAV78448
20	38	51.4	36	18	AAV31508
21	38	51.4	215	20	AAV37046
22	38	51.4	293	20	AAV25095
23	38	51.4	486	20	AAV24427
24	38	51.4	532	20	AAV24425
25	38	51.4	578	20	AAV24423
26	37	50.0	263	10	AAV92003
27	37	50.0	263	20	AAV29859
28	37	50.0	263	21	AAV13014
29	37	50.0	263	22	AAV48846
30	37	50.0	286	20	AAV25094
31	37	50.0	286	21	AAV70235
32	37	50.0	345	19	AAV46493
33	37	50.0	384	14	AAV34649
34	37	50.0	384	18	AAV28545
35	37	50.0	384	21	AAV70159
36	37	50.0	384	21	AAV70174
37	37	50.0	384	21	AAV70175
38	37	50.0	384	21	AAV70176
39	37	50.0	384	21	AAV70177
40	37	50.0	384	21	AAV70178
41	37	50.0	384	21	AAV70179
42	37	50.0	384	21	AAV70180
43	37	50.0	384	21	AAV70181
44	37	50.0	384	21	AAV70182
45	37	50.0	384	21	AAV70183

ALIGNMENTS

RESULT 1	
AAV81995	AAV81995 standard; peptide: 12 AA.
ID	AAV81995;
AC	AAV81995;
XX	16-OCT-2000 (first entry)
DT	
DE	Human high molecular weight kininogen domain 5 fragment #4.
XX	
KW	Human: high molecular weight kininogen; HK;
KW	two-chain high molecular weight kininogen; HKa;
KW	angiogenesis inhibition; tumour; cancer; ocular disorder;
KW	rheumatoid arthritis; endothelial cell apoptosis.
OS	Homo sapiens.
XX	
PV	WO200027866-A1.
XX	
PD	18-MAY-2000.
XX	
PF	05-NOV-1999; 99WO-US26419.
XX	
PR	10-NOV-1998; 98US-0107833.
XX	
PA	(UTEM) UNIV TEMPLE.
PA	(MCCR/) MCCR R K.
XX	
PI	McCrack RK.
XX	
DR	WPI: 2000-376483/32.
XX	
PT	A pharmaceutical composition used to inhibit angiogenesis, inhibit endothelial cell proliferation, and induce endothelial cell apoptosis

PT -
XX Claim 10; Page 29; 52pp; English.
PS
XX
CC The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain
CC high molecular weight kininogen (Hka) by plasma kallikrein. Hka or a
CC synthetic compound comprising part or all of the present sequence may
CC be used in a pharmaceutical composition for inhibiting angiogenesis.
CC Angiogenesis occurs in a number of disease states, such as tumour
CC formation and expansion, and certain ocular disorders. It can also
CC occur in a Rheumatoid joint, hastening joint destruction by allowing
CC an influx of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial
CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.
XX
SQ Sequence 12 AA;
XX
OY Query Match 100.0%; Score 74; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GKKNGKHNGMKT 12
1 gkknghngmkt 12
Db
OY 1 GKKNGKHNGMKT 12
1 gkknghngmkt 12
Db
RESULT 2
AAB06337
ID AAB06337 standard; Protein; 16 AA.
XX
AC AAB06337;
XX
DT 16-OCT-2000 (first entry)
XX
DE Human two-chain high molecular weight kininogen domain 5 fragment #9.
XX
XX Human; high molecular weight kininogen; HK;
XX two-chain high molecular weight kininogen; Hka;
XX angiogenesis inhibition; tumour; cancer; ocular disorder;
XX rheumatoid arthritis; endothelial cell apoptosis.
XX
OS Homo sapiens.
XX
PN WO200027866-A1.
XX
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26419.
XX
PR 10-NOV-1998; 98US-0107833.
XX
XX (UTEM) UNIV TEMPLE.
XX (MCCR/) MCCRAE R K.
XX
XX McCrae RK;
XX
DR WPI; 2000-376483/32.
XX
PT A pharmaceutical composition used to inhibit angiogenesis, inhibit
PT endothelial cell proliferation, and induce endothelial cell apoptosis
PT
XX
XX
PS Claim 15; Page 29; 52pp; English.
XX
CC The present sequence is derived from human two-chain high molecular
CC weight kininogen (Hka) domain 5. Hka is product of high molecular
CC weight kininogen (HK) cleavage by plasma kallikrein. HK is a 120 kD

CC glycoprotein which binds with high affinity to endothelial cells. Hka
CC or a synthetic compound comprising the present sequence may be used in
CC a pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
CC occurs in a number of disease states, such as tumour formation and
CC expansion, and certain ocular disorders. It can also occur in a
CC Rheumatoid joint, hastening joint destruction by allowing an influx
CC of leukocytes. The composition may inhibit angiogenesis by inhibiting
CC endothelial cell proliferation or by inducing endothelial cell
CC apoptosis. Peptides used in the composition may be recombinant peptides,
CC natural peptides, or synthetic peptides. They may also be chemically
CC synthesised, using, for example, solid phase synthesis methods.
XX
SQ Sequence 16 AA;
XX
OY Query Match 100.0%; Score 74; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GKKNGKHNGMKT 12
5 gkknghngmkt 16
Db
OY 1 GKKNGKHNGMKT 12
5 gkknghngmkt 16
Db
RESULT 3
AAY81997
ID AAY81997 standard; Peptide; 28 AA.
XX
AC AAY81997;
XX
DT 16-OCT-2000 (first entry)
XX
DE Human high molecular weight kininogen domain 5 fragment #6.
XX
XX Human; high molecular weight kininogen; HK;
XX two-chain high molecular weight kininogen; Hka;
XX angiogenesis inhibition; tumour; cancer; ocular disorder;
XX rheumatoid arthritis; endothelial cell apoptosis.
XX
OS Homo sapiens.
XX
PN WO200027866-A1.
XX
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26419.
XX
PR 10-NOV-1998; 98US-0107833.
XX
XX (UTEM) UNIV TEMPLE.
XX (MCCR/) MCCRAE R K.
XX
XX McCrae RK;
XX
DR WPI; 2000-376483/32.
XX
PT A pharmaceutical composition used to inhibit angiogenesis, inhibit
PT endothelial cell proliferation, and induce endothelial cell apoptosis
PT
XX
XX
PS Claim 13; Page 29; 52pp; English.
XX
CC The present sequence is derived from human high molecular weight
CC kininogen (HK) domain 5. HK is a 120 kD glycoprotein which binds with
CC high affinity to endothelial cells, where it is cleaved to two-chain
CC high molecular weight kininogen (Hka) by plasma kallikrein. Hka or a
CC synthetic compound comprising the present sequence may be used in a
CC pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
CC occurs in a number of disease states, such as tumour formation and
CC expansion, and certain ocular disorders. It can also occur in a
CC Rheumatoid joint, hastening joint destruction by allowing an influx
CC of leukocytes. The composition may inhibit angiogenesis by
CC inhibiting endothelial cell proliferation or by inducing endothelial

CC cell apoptosis. Peptides used in the composition may be recombinant
CC peptides, natural peptides, or synthetic peptides. They may also be
CC chemically synthesised, using, for example, solid phase synthesis
CC methods.

SO Sequence 28 AA;

Query Match 100.0%; Score 74; DB 21; Length 28;

Best Local Similarity 100.0%; Pred. No. 3.1e-05; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKNGKHNGWKT 12
|
17 gkngkhngwkt 28

RESULT 4

AA93351
ID AAY93351 standard; peptide: 94 AA.

XX AAY93351;

XX 04-SEP-2000 (first entry)

XX Light chain of human high molecular weight kininogen analogue.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;

XX plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

XX endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

XX Homo sapiens.

XX WO200027415-A2.

XX 18-MAY-2000.

XX 09-NOV-1999; 99WO-US26377.

XX 10-NOV-1998; 98US-0107844.

XX (UTEM) UNIV TEMPLE.

XX (DUPO) DUPONT PHARM CO.

XX (COLM/) COLMAN W R.

XX (MOSA/) MOUSA A S.

XX Colman WR, Mousa AS;

XX WPI; 2000-376306/32.

XX Method for inhibiting endothelial cell proliferation, using compound

XX that inhibit endothelial cell migration

XX Claim 8; Page 39; 41pp; English.

XX The present sequence represents an analogue of the light chain of human

XX high molecular weight kininogen. High molecular weight kininogen is a

XX 120 kDa glycoprotein which binds with high affinity to endothelial cells,

XX where it is cleaved by plasma kallikrein into heavy and light chains.

XX Analogues of high molecular weight kininogen are used in the method

XX of the invention. The specification describes a method of inhibiting

XX endothelial cell proliferation. The method comprises contacting

XX endothelial cells with a compound containing high molecular weight

XX kininogen analogues. The method and the compounds can be used for

XX inhibiting endothelial cell proliferation. The compounds can also be

XX used for inhibiting angiogenesis. The compounds can also be used to

XX inhibit migration of endothelial cells to vitronectin.

Best Local Similarity 100.0%; Pred. No. 0.00011;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKNGKHNGWKT 12
|
73 gkngkhngwkt 84

RESULT 5

AA93353
ID AAY93353 standard; peptide: 179 AA.

XX AAY93353;

XX 04-SEP-2000 (first entry)

XX Light chain of human high molecular weight kininogen analogue.

XX Human; high molecular weight kininogen; glycoprotein; endothelial cell;

XX plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;

XX endothelial cell proliferation; endothelial cell migration; vitronectin.

XX Synthetic.

XX Homo sapiens.

XX WO200027415-A2.

XX 18-MAY-2000.

XX 09-NOV-1999; 99WO-US26377.

XX 10-NOV-1998; 98US-0107844.

XX (UTEM) UNIV TEMPLE.

XX (DUPO) DUPONT PHARM CO.

XX (COLM/) COLMAN W R.

XX (MOSA/) MOUSA A S.

XX Colman WR, Mousa AS;

XX WPI; 2000-376306/32.

XX Method for inhibiting endothelial cell proliferation, using compound

XX that inhibit endothelial cell migration

XX Claim 11; Page 40-41; 41pp; English.

XX The present sequence represents an analogue of the light chain of human

XX high molecular weight kininogen. High molecular weight kininogen is a

XX 120 kDa glycoprotein which binds with high affinity to endothelial cells,

XX where it is cleaved by plasma kallikrein into heavy and light chains.

XX Analogues of high molecular weight kininogen are used in the method

XX of the invention. The specification describes a method of inhibiting

XX endothelial cell proliferation. The method comprises contacting

XX endothelial cells with a compound containing high molecular weight

XX kininogen analogues. The method and the compounds can be used for

XX inhibiting endothelial cell proliferation. The compounds can also be

XX used for inhibiting angiogenesis. The compounds can also be used to

XX inhibit migration of endothelial cells to vitronectin.

SO Sequence 179 AA;

Query Match 100.0%; Score 74; DB 21; Length 179;

Best Local Similarity 100.0%; Pred. No. 0.00021;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKNGKHNGWKT 12
|
45 gkngkhngwkt 56

RESULT 6

Query Match 100.0%; Score 74; DB 21; Length 94;

AAV93349
ID AAV93349 standard; peptide; 186 AA.
XX
AC AAV93349;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen analogue.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
PM WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOSA/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
XX
DR WPI: 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX
PS Claim 9; Page 38; 41pp; English.
XX
CC The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 186 AA:

Query Match 100.0%; Score 74; DB 21; Length 186;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKKNGKHNGWKT 12
|
Db 52 gKngkhngwkt 63

RESULT 7
AAV93342
ID AAV93342 standard; protein; 255 AA.
XX
AC AAV93342;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen.
XX

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Homo sapiens.
XX
PM WO200027415-A2.
XX
PD 18-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26377.
XX
PR 10-NOV-1998; 98US-0107844.
XX
PA (UTEM) UNIV TEMPLE.
PA (DUPO) DUPONT PHARM CO.
PA (COLM/) COLMAN W R.
PA (MOSA/) MOUSA A S.
XX
PI Colman WR, Mousa AS;
XX
DR WPI: 2000-376306/32.
XX
PT Method for inhibiting endothelial cell proliferation, using compound
PT that inhibit endothelial cell migration
XX
PS Disclosure; Page 3; 41pp; English.
XX
CC The present sequence represents the light chain of human high molecular
CC weight kininogen. High molecular weight kininogen is a 120 kDa
CC glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.
XX
SQ Sequence 255 AA:

Query Match 100.0%; Score 74; DB 21; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GKKNGKHNGWKT 12
|
Db 121 gKngkhngwkt 132

RESULT 8
AAV93345
ID AAV93345 standard; peptide; 47 AA.
XX
AC AAV93345;
XX
DT 04-SEP-2000 (first entry)
XX
DE Light chain of human high molecular weight kininogen fragment.
XX
KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.
XX
OS Synthetic.
OS Homo sapiens.
XX
PM WO200027415-A2.
XX
PD 18-MAY-2000.

CC on fibroblasts.
XX
SQ Sequence 63 AA:

Query Match 93.2%; Score 69; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GKKGKHNKWK 11
| | | | | | | | | | | | |
Db 53 gkknghkngwk 63

RESULT 11
AA93347
ID AAY93347 standard; peptide: 83 AA.

AC AAY93347;

DT 04-SEP-2000 (first entry)

DE Light chain of human high molecular weight kininogen analogue.

KW Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW endothelial cell proliferation; endothelial cell migration; vitronectin.

OS Synthetic.
OS Homo sapiens.

PN WO200027415-A2.

PD 18-MAY-2000.

PF 09-NOV-1999; 99WO-US26377.

PR 10-NOV-1998; 98US-0107844.

PA (UTEM) UNIV TEMPLE.

PA (DUPON) DUPONT PHARM CO.

PA (COLM/) COLMAN W R.

PA (MOUS/) MOUSA A S.

PI Colman WR, Mousa AS;

PI WPI; 2000-376306/32.

PT Method for inhibiting endothelial cell proliferation, using compound

PT that inhibit endothelial cell migration -

PS Claim 5; Page 37; 41pp; English.

XX The present sequence represents an analogue of the light chain of human
CC high molecular weight kininogen. High molecular weight kininogen is a
CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
CC where it is cleaved by plasma kallikrein into heavy and light chains.
CC Analogues of high molecular weight kininogen are used in the method
CC of the invention. The specification describes a method of inhibiting
CC endothelial cell proliferation. The method comprises contacting
CC endothelial cells with a compound containing high molecular weight
CC kininogen analogues. The method and the compounds can be used for
CC inhibiting endothelial cell proliferation. The compounds can also be
CC used for inhibiting angiogenesis. The compounds can also be used to
CC inhibit migration of endothelial cells to vitronectin.

SQ Sequence 83 AA:

Query Match 93.2%; Score 69; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 0.00058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKGKHNKWK 11
| | | | | | | | | | | | |
Db 73 gkknghkngwk 83

RESULT 12
AAR75181
ID AAR75181 standard; peptide: 131 AA.

AC AAR75181;

DT 05-DEC-1995 (first entry)

DE Partial peptide of human HMW kininogen fragment 1.2.

KW high molecular weight; kininogen; fragment; 1.2; 1: 2; partial;
KW wound treating agent; human; growth promotion; fibroblast.

OS Homo sapiens.

PN JP07082172-A.

PD 28-MAR-1995;

PF 17-SEP-1993; 93JP-0230616.

PR 17-SEP-1993; 93JP-0230616.

PA (FARH) HOECHST JAPAN KK.

PA WPI; 1995-158909/21.

PT A wound treating agent contg. a partial peptide of kininogen -

PT have growth promotion activity of fibroblasts.

PS Claim 7; Page 7; 8pp; Japanese.

XX AAR75181 is a partial peptide corresponding to human kininogen
CC fragment 1.2, amino acids 390-520. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compsns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.

SQ Sequence 131 AA:

Query Match 93.2%; Score 69; DB 16; Length 131;
Best Local Similarity 100.0%; Pred. No. 0.00093;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKGKHNKWK 11
| | | | | | | | | | | | |

Db 121 gkknghkngwk 131

RESULT 13

AAR75180
ID AAR75180 standard; peptide: 41 AA.

AC AAR75180;

DT 05-DEC-1995 (first entry)

DE Partial peptide of HMW kininogen fragment 2.

KW high molecular weight; kininogen; fragment; 1.2; 1: 2; partial;
KW wound treating agent; bovine; growth promotion; fibroblast.

OS Bos taurus.

PN JP07082172-A.

PD 28-MAR-1995.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX (FARH) HOECHST JAPAN KK.
XX
XX WPI: 1995-158909/21.
XX
XX A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
XX
XX Claim 6; Page 7; 8pp; Japanese.
XX
XX AAR75179 is a partial peptide corresponding to bovine kininogen
CC fragment 2, amino acids 456-496. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compsns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
XX Sequence 41 AA:
SQ
Query Match 62.2%; Score 46; DB 16; Length 41;
Best Local Similarity 63.6%; Pred. No. 1.1;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 1 GKKNGKHNGWK 11
| | | | | |
Db 31 gknngkhydw 41
RESULT 14
AAR75178
ID AAR75178 standard; peptide; 110 AA.
XX
XX AAR75178;
XX
XX 05-DEC-1995 (first entry)
XX
XX Partial peptide of HMW kininogen fragment 1.2.
XX
XX high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KM wound treating agent; bovine; growth promotion; fibroblast.
XX
XX Bos taurus.
XX
XX Key Location/Qualifiers
FH Misc-difference 12 /label- Pro, Thr
FT Misc-difference 15 /label- Val or Leu
FT Misc-difference 69 /label- Lys or Arg
FT
XX JPO7082172-A.
XX
XX 28-MAR-1995.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX 17-SEP-1993; 93JP-0230616.
XX
XX (FARH) HOECHST JAPAN KK.
XX
XX WPI: 1995-158909/21.
XX
XX A wound treating agent contg. a partial peptide of kininogen -
PT have growth promotion activity of fibroblasts.
XX
XX Claim 4; Page 6; 8pp; Japanese.

XX
XX AAR75178 is a partial peptide corresponding to bovine kininogen
CC fragment 1.2, amino acids 387-496. Partial peptides of bovine and
CC human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC agent compsns. and act as the active component. The fragments are
CC useful in wound treating because they have growth promotion activity
CC on fibroblasts.
XX
XX Sequence 110 AA:
SQ
Query Match 62.2%; Score 46; DB 16; Length 110;
Best Local Similarity 63.6%; Pred. No. 3.1;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 1 GKKNGKHNGWK 11
| | | | | |
Db 100 gknngkhydw 110
RESULT 15
AAY71879
ID AAY71879 standard; peptide; 19 AA.
XX
XX AAY71879;
XX
XX 26-MAR-2001 (first entry)
XX
XX Human HKH20 peptide derived from domain 5 of H-Kininogen (479-498 aa).
XX
XX Human; heparin binding protein; HBP; antiinflammatory; cardiovascular;
KW immunosuppressive; vasotropic; prevention; treatment; bradykinin;
KW aprotinin; H-kininogen; HK; systemic inflammatory response syndrome;
KW pre-kallikrein; ischaemia reperfusion; anaphylaxis; allograft rejection;
KW adult respiratory distress syndrome; HKH20 peptide.
XX
XX Homo sapiens.
XX
XX WO2000066151-A1.
XX
XX 09-NOV-2000.
XX
XX 28-APR-2000; 2000WO-DK00213.
XX
XX 29-APR-1999; 99US-0132748.
PR 06-MAY-1999; 99DK-0000613.
PR 01-OCT-1999; 99DK-0001402.
PR 01-OCT-1999; 99US-0157384.
XX
XX (NOVO) NOVO NORDISK AS.
XX
XX Flodgaard HJ, Lindom L, Bjorn S;
XX
XX WPI: 2000-687445/67.
XX
XX Treating systemic inflammatory response syndrome, ischaemia reperfusion,
PT anaphylaxis and allograft rejection by modulating release of bradykinin
PT
XX
XX Example 2; Page 39; 75pp; English.
XX
XX The patent discloses a method for preventing or treating a disorder
CC resulting from the release of bradykinin in a mammal which produces
CC a heparin-binding protein (HBP) that binds to a HBP antagonist. This
CC method involves administration of a mammalian HBP antagonist (especially
CC aprotinin) and/or monoclonal antibodies that bind to prekallikrein-
CC H-kininogen complexes in the HBP, to decrease the release of bradykinin
CC in the mammal. The antagonists of HBP (e.g. aprotinin) decrease the
CC permeability of the endothelial cells and are used to prevent or treat
CC disorders resulting from the release of bradykinin such as systemic
CC inflammatory response syndrome, ischaemia reperfusion, anaphylaxis
CC and/or allograft rejection. They are also used to treat adult
CC respiratory distress syndrome.

CC The present sequence is HKH20 peptide which is derived from the
 CC domain 5 of human H-Kininogen (HK) protein (479-498 residues).
 CC HKH20 treatment of endothelial cells inhibits or prevents the HBP-
 CC induced increase in permeability of the endothelial cells.
 XX

SQ Sequence 19 AA;

Query Match 55.4%; Score 41; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKKNGKH 7
 |||||
 Db 13 gkkngh 19

Search completed: July 6, 2001, 09:09:17
 Job time: 123 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:21 ; Search time 56.74 Seconds
(without alignments)
4.260 Million cell updates/sec

Title: US-09-437-912-4
Perfect score: 74
Sequence: 1 GKKNGKNGWKT 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_Aa.*
1: /cgn2-6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2-6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2-6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2-6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2-6/ptodata/2/1aa/PTCTUS_COMB.pep.*
6: /cgn2-6/ptodata/2/1aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	54.1	376	1	US-08-594-031-100 Sequence 100, App
2	40	54.1	376	1	US-08-594-031-102 Sequence 102, App
3	39	52.7	60	2	US-08-117-952-787 Sequence 787, App
4	39	52.7	60	2	US-08-117-952-788 Sequence 788, App
5	39	52.7	3177	2	US-08-477-451-4 Sequence 4, Appli
6	38	51.4	293	3	US-09-203-716-2 Sequence 2, Appli
7	37	50.0	126	6	5514582-43 Patent No. 5514582
8	37	50.0	263	1	US-07-906-983-2 Sequence 2, Appli
9	37	50.0	286	3	US-09-203-716-1 Sequence 1, Appli
10	37	50.0	345	1	US-08-183-214-12 Sequence 12, Appli
11	37	50.0	384	1	US-07-783-706-2 Sequence 2, Appli
12	37	50.0	384	4	US-08-445-342A-2 Sequence 2, Appli
13	37	50.0	384	4	US-09-066-481-2 Sequence 2, Appli
14	37	50.0	384	5	PCT-US92-09124-2 Sequence 2, Appli
15	37	50.0	590	1	US-08-221-817-14 Sequence 14, Appli
16	37	50.0	590	1	US-08-454-439-14 Sequence 14, Appli
17	37	50.0	590	5	PCT-US94-10487-14 Sequence 14, Appli
18	37	50.0	642	1	US-08-414-926A-25 Sequence 25, Appli
19	37	50.0	642	2	US-08-926-922-25 Sequence 25, Appli
20	37	50.0	642	3	US-09-253-682-25 Sequence 25, Appli
21	36	48.6	8	1	US-08-690-090A-8 Sequence 8, Appli
22	36	48.6	29	4	US-08-816-346-44 Sequence 44, Appli
23	36	48.6	29	4	US-09-335-411-44 Sequence 44, Appli
24	36	48.6	51	4	US-08-816-346-40 Sequence 40, Appli
25	36	48.6	51	4	US-09-335-411-40 Sequence 4, Appli
26	36	48.6	230	4	US-08-768-373-4 Sequence 4, Appli
27	36	48.6	233	2	US-08-458-568A-4 Sequence 4, Appli

28	36	48.6	951	4	US-08-816-346-58 Sequence 58, Appli
29	36	48.6	951	4	US-09-335-411-58 Sequence 58, Appli
30	36	48.6	952	2	US-08-788-674-5 Sequence 5, Appli
31	36	48.6	952	4	US-08-816-346-4 Sequence 4, Appli
32	36	48.6	952	4	US-09-335-411-4 Sequence 4, Appli
33	36	48.6	369	4	US-08-061-376-5 Sequence 5, Appli
34	35.5	48.0	367	2	US-08-655-704B-17 Sequence 17, Appli
35	35.5	48.0	367	2	US-09-107-755-17 Sequence 17, Appli
36	35.5	48.0	1507	6	5268270-2 Patent No. 5268270
37	35	47.3	155	3	US-09-203-716-4 Sequence 4, Appli
38	35	47.3	615	4	US-08-989-299-11 Sequence 11, Appli
39	35	47.3	3898	4	US-08-750-717-2 Sequence 2, Appli
40	34	45.9	32	1	US-08-190-802A-69 Sequence 69, Appli
41	34	45.9	32	1	US-08-190-802A-128 Sequence 128, App
42	34	45.9	32	1	US-08-190-802A-172 Sequence 172, App
43	34	45.9	212	1	US-08-462-965A-2 Sequence 2, Appli
44	34	45.9	212	1	US-08-462-169B-21 Sequence 21, Appli
45	34	45.9	212	3	US-09-103-079-21 Sequence 21, Appli

ALIGNMENTS

RESULT 1
US-08-594-031-100
: Sequence 100, Application US/08594031
: Patent No. 5783182
: GENERAL INFORMATION:
: APPLICANT: THOMPSON, Timothy C.
: TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
: NUMBER OF SEQUENCES: 175
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BAKER & BOTTS, L.L.P.
: STREET: 1299 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20004-2400
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/594,031
: FILING DATE: 30-JAN-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/006,838
: FILING DATE: 16-NOV-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Remenick, James
: REGISTRATION NUMBER: 36,902
: REFERENCE/DOCKET NUMBER: 0A146-0110
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-639-7700
: TELEFAX: 202-639-7890
: TELEX:
: INFORMATION FOR SEQ ID NO: 100:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 376 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: US-08-594-031-100
Query Match 54.1%, Score 40; DB 1; Length 376;

Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 4 NGKHGKWK 11
:||||:
Db 114 GKHGKWK 121

RESULT 2
US-08-594-031-102
; Sequence 102, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; US-08-594-031-102

Query Match 54.1%; Score 40; DB 1; Length 376;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 4 NGKHGKWK 11
:||||:
Db 114 GKHGKWK 121

RESULT 3
US-08-117-952-787
; Sequence 787, Application US/08117952
; Patent No. 5851760
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Smith, Michael W.

;; TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
;; TITLE OF INVENTION: SAMPLED MAPS OF COMPLEX GENOMES
;; NUMBER OF SEQUENCES: 797
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
;; STREET: 444 South Flower Street, Suite 2000
;; CITY: Los Angeles
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 90071
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/117,952
;; FILING DATE: 07-SEP-1993
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/078,471
;; FILING DATE: 15-JUN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Reiter, Stephen E.
;; REGISTRATION NUMBER: 31,192
;; REFERENCE/DOCKET NUMBER: P41 9423
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619-546-4737
;; TELEFAX: 619-546-9392
;; INFORMATION FOR SEQ ID NO: 787:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 60 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: Internal
;; US-08-117-952-787

Query Match 52.7%; Score 39; DB 2; Length 60;
Best Local Similarity 53.3%; Pred. No. 7;
Matches 8; Conservative 0; Mismatches 3; Indels 4; Gaps 1;
QY 1 GKXGKHNG---WK 11
||||| | | | |
Db 10 GKXSPHECKRIMWK 24

RESULT 4
US-08-117-952-788
; Sequence 788, Application US/08117952
; Patent No. 5851760
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Smith, Michael W.
; TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
; TITLE OF INVENTION: SAMPLED MAPS OF COMPLEX GENOMES
; NUMBER OF SEQUENCES: 797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,952
; FILING DATE: 07-SEP-1993

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/078,471
FILING DATE: 15-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9423
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 788:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-117-952-788

Query Match 52.7%; Score 39; DB 2; Length 60;
Best Local Similarity 53.3%; Pred. No. 7;
Matches 8; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

OY 1 GKNGKNGHNG---WK 11
||| | | | |
Db 10 GKNGSPHCKGLIWK 24

RESULT 5
US-08-477-451-4
Sequence 4, Application US/08477451
Patent No. 5928865
GENERAL INFORMATION:
APPLICANT: Covacci, Antonello
TITLE OF INVENTION: Helicobacter Pylori CagI Region
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,451
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0335.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2708
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3177 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-451-4

Query Match 52.7%; Score 39; DB 2; Length 3177;
Best Local Similarity 60.0%; Pred. No. 4.3e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 2 KNGKNGKNGWK 11
||| :||:|
Db 910 KKNRNYNGYK 919

RESULT 6
US-09-203-716-2
Sequence 2, Application US/09203716
Patent No. 6001653
GENERAL INFORMATION:
APPLICANT: Crooke, Stanley T.
APPLICANT: Lima, Walter F.
APPLICANT: Wu, Hongjiang
TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
FILE REFERENCE: ISPh-0333
CURRENT APPLICATION NUMBER: US/09/203,716
CURRENT FILING DATE: 1998-12-02
EARLIER APPLICATION NUMBER: 60/067,458
EARLIER FILING DATE: 1997-12-04
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 293
TYPE: PRT
ORGANISM: Gallus sp.
US-09-203-716-2

Query Match 51.4%; Score 38; DB 3; Length 293;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 KNGKNGKNGWK 12
:| | |||:|
Db 230 ENMKGTNGWRT 239

RESULT 7
5514582-43
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 43
LENGTH: 126
5514582-43

Query Match 50.0%; Score 37; DB 6; Length 126;
Best Local Similarity 62.5%; Pred. No. 32;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 NGKNGKNGWK 11
||:||||:
Db 75 NGRHNGYE 82

RESULT 8

US-07-906-983-2
; Sequence 2, Application US/07906983
; Patent No. 5187268
; GENERAL INFORMATION:
; APPLICANT: Kotwal, Girish
; APPLICANT: Moss, Bernard
; TITLE OF INVENTION: Synthetic, Anti-Complement Protein and
; TITLE OF INVENTION: the Gene Encoding Same
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,983
; FILING DATE: 1992/07/01
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-983-2

Query Match 50.0%; Score 37; DB 1; Length 263;
Best Local Similarity 62.5%; Pred. No. 68;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 KNGKNGMK 11
|:||||:
Db 156 NGRHNGYE 163

RESULT 9
US-09-203-716-1
; Sequence 1, Application US/09203716
; Patent No. 6001653
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Lima, Walter F.
; APPLICANT: Wu, Hongjiang
; TITLE OF INVENTION: Human RNase H Compositions and Uses Thereof
; FILE REFERENCE: ISFH-0333
; CURRENT APPLICATION NUMBER: US/09/203,716
; EARLIER FILING DATE: 1998-12-02
; EARLIER APPLICATION NUMBER: 60/067,458
; EARLIER FILING DATE: 1997-12-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-203-716-1

Query Match 50.0%; Score 37; DB 3; Length 286;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 KNGKMT 12
|:||||:
Db 226 KNGKMT 232

RESULT 10
US-08-183-214-12
; Sequence 12, Application US/08183214
; Patent No. 5716816
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Stanley, Sally J.
; APPLICANT: Nightingale, Maria S.
; APPLICANT: Murlagh, Jr., James J.
; APPLICANT: Monaco, Lucia
; APPLICANT: Takada, Tatsuyuki
; TITLE OF INVENTION: CLONES ENCODING MAMMALIAN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/183,214
; FILING DATE: 14-JAN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,231
; FILING DATE: 22-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-5043
; TELEFAX: 415-543-9600
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-183-214-12

Query Match 50.0%; Score 37; DB 1; Length 345;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KNGKNGMK 11
|:||||:
Db 99 KPGKPNQWR 107

RESULT 11
US-07-783-706-2
; Sequence 2, Application US/07783706
; Patent No. 5714376
; GENERAL INFORMATION:
; APPLICANT: Sasisekharan, Ramnath

APPLICANT: Moreman, Kelley
APPLICANT: Cooney, Charles L.
APPLICANT: Langer, Robert S.
TITLE OF INVENTION: The Heparinase Gene from Flavobacterium
TITLE OF INVENTION: Heparinum
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street, Suite 3100
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/783,706
FILING DATE: 19911023
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MITS546
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6535
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: YES
ANTI-SENSE: YES
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Flavobacterium heparinum
US-07-783-706-2

Query Match 50.0%; Score 37; DB 1; Length 384;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 GKHNGWK 11
DB 220 GKPNGWK 226

RESULT 12
US-08-445-342A-2
Sequence 2, Application US/08445342A
Patent No. 5830726
GENERAL INFORMATION:
APPLICANT: Sasisekharan, Rammath
APPLICANT: Moreman, Kelley
APPLICANT: Cooney, Charles L.
APPLICANT: Zimmerman, Joseph, J.
APPLICANT: Langer, Robert, S.
TITLE OF INVENTION: The Heparinase gene from Flavobacterium
TITLE OF INVENTION: Heparinum
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,342A
FILING DATE: 19-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,706
FILING DATE: 23-OCT-1991
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MIT 5546 div
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-445-342A-2

Query Match 50.0%; Score 37; DB 2; Length 384;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 GKHNGWK 11
DB 220 GKPNGWK 226

RESULT 13
US-09-066-481-2
Sequence 2, Application US/09066481B
Patent No. 6217863
GENERAL INFORMATION:
APPLICANT: GODAVARTI, RANGANATHAN
APPLICANT: SASISEKHARAN, RAMMATH
APPLICANT: ERNST, STEFFAN
APPLICANT: GANESH VENKATARAMAN
APPLICANT: COONEY, CHARLES L.
TITLE OF INVENTION: RATIONALLY DESIGNED POLYSACCHARIDE
TITLE OF INVENTION: LYASES DERIVED FROM HEPARINASE I
FILE REFERENCE: M0656/7038/HCL
CURRENT APPLICATION NUMBER: US/09/066,481B
CURRENT FILING DATE: 1999-01-19
EARLIER APPLICATION NUMBER: US 60/008,069
EARLIER FILING DATE: 1995-10-30
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Flavobacterium Heparinum
US-09-066-481-2

Query Match 50.0%; Score 37; DB 4; Length 384;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 GKHNGWK 11
DB 220 GKPNGWK 226

RESULT 14
PCT-US92-09124-2
Sequence 2, Application PC/TUS9209124
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute, of Technology
TITLE OF INVENTION: The Heparinase Gene from Flavobacterium
TITLE OF INVENTION: Heparinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09124
FILING DATE: 19921022
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: M15546
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6535
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Flavobacterium heparinum
PCT-US92-09124-2

Query Match 50.0%; Score 37; DB 5; Length 384;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 GRKNGWK 11
|||
Db 220 GRKNGWK 226

RESULT 15
US-08-221-817-14
Sequence 14, Application US/08221817
Patent No. 5532151
GENERAL INFORMATION:
APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hoekstra, Merle F.
TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois

COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,817
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5532151and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 590 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-221-817-14

Query Match 50.0%; Score 37; DB 1; Length 590;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 GRKNGKNGWK 11
|||
Db 21 GRKNGKNGWK 31

Search completed: July 6, 2001, 09:10:22
Job time: 188 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:17:59 ; Search time 73.59 Seconds

(without alignments)
12.421 Million cell updates/sec

Title: US-09-437-912-4

Sequence: 1 GKNKGKHNKMK 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	644	1 KGHUHI	kininogen, HMW pre
2	51	68.9	619	1 KGBOH2	kininogen, HMW II
3	51	68.9	621	1 KGBOH1	kininogen, HMW I P
4	47	63.5	275	2 T35064	probable integral
5	46	62.2	169	2 G71333	probable ribonucle
6	43	58.1	516	2 T49422	RAD57 related prot
7	42	56.8	374	2 T11662	probable transcript
8	42	56.8	1009	2 T31081	ccs3 protein - rat
9	41	55.4	145	2 F81807	ribonuclease H (EC
10	41	55.4	145	2 H81061	ribonuclease H (EC
11	41	55.4	344	2 S34153	msl101-1 protein -
12	41	55.4	385	2 S72275	nuclease taxc - Esc
13	41	55.4	552	2 D81290	probable capsule p
14	41	55.4	857	2 T05352	hypothetical prote
15	40	54.1	198	2 T27150	hypothetical prote
16	40	54.1	721	2 T40945	hypothetical prote
17	40	54.1	1430	2 T27924	hypothetical prote
18	39	52.7	304	2 G70300	ribosomal protein
19	39	52.7	311	2 E86746	cell division prot
20	39	52.7	616	2 U01441	hypothetical 67K p
21	39	52.7	789	2 S46631	acornate hydratase
22	39	52.7	874	4 GNMUER	retrovirus-related
23	39	52.7	2042	2 T18399	variant-specific s
24	38.5	53.0	202	2 H86689	prophage ps2 prote
25	38.5	53.0	491	2 A49993	glycylpeptide N-te
26	38	51.4	60	2 UC2127	G protein-coupled
27	38	51.4	81	2 A86772	hypothetical prote
28	38	51.4	133	2 T52527	probable ribonucle
29	38	51.4	179	2 G75462	ribonuclease H - D

30	38	51.4	221	2 T45044	hypothetical prote
31	38	51.4	235	2 T35324	probable ribonucle
32	38	51.4	861	2 B84963	DNA topoisomerase
33	38	51.4	1534	2 S59604	DNA (cytosine-5-)-
34	37.5	50.7	299	2 T47917	probable transcript
35	37.5	50.7	359	2 A86804	prophage p13 prote
36	37.5	50.7	891	2 T40137	hypothetical serin
37	37	50.0	102	1 S78234	ribosomal protein
38	37	50.0	174	2 S73219	ribosomal protein
39	37	50.0	194	2 T16556	hypothetical prote
40	37	50.0	224	2 S71749	DCI protein precu
41	37	50.0	263	1 MMVZSP	apolipoprotein H h
42	37	50.0	298	2 A84100	cell-division prot
43	37	50.0	312	2 T37922	hypothetical prote
44	37	50.0	357	2 B47411	Adribosylarginine
45	37	50.0	383	2 T21453	hypothetical prote

ALIGNMENTS

RESULT 1
KGHOHI
kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-chiol proteinase inhibitor; preprokininogen; prokininogen
M:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030;
R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shikawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A>Title: Isolation of a human cDNA for alpha-2-chiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:85122621
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:K02566; NID:G177889
R:Auerwald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A>Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:93223854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 1-592, 'I', 594-644 <PAK>
A:Cross-references: GB:M1437; NID:G186751; PID:NAB59550.1; PID:G386852
R:Auerwald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A>Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:93223854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: 'ANSM', 235-377 <AUE>
A>Note: differences are due to known cloning artifacts
R:Loetsch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A>Title: The amino acid sequence of the light chain of human high-molecular-mass kin
A:Reference number: A91153; MUID:86030270
A:Accession: A91153
A:Molecule type: protein
A:Residues: 379-644 <LOT>
A>Note: the bradykinin sequence preceding the light chain sequence was not determined
R:Kellermann, J.; Loetsch, F.; Henschen, A.; Muller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A>Title: Completion of the primary structure of human high-molecular-mass kininogen.
A:Reference number: A24871; MUID:86108361
A:Accession: A24871
A:Molecule type: protein
A:Residues: 'Z', 20-380 <KEI>
R:Kellermann, J.; Loetsch, F.; Henschen, A.; Muller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
A>Title: Amino acid sequence of the light chain of human high molecular mass kininoge
A:Reference number: A27899
A:Accession: A27899

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <REL2>
 R:Minidrolu, T.; Carretero, O.A.; Proud, D.; Seclit, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88209021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A31905; MUID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAE>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plas
 A:Reference number: A34030; MUID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Lenarctic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
 A:Reference number: S02482; MUID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19, 169-192, 310-314, 381-389 <LEN1>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Lenarctic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55239; MUID:95251593
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Straczek, J.; Maschl, F.; Le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; MUID:96033974
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evolut
 A:Reference number: A92545; MUID:85234583
 A:Contents: annotation; gene organization

R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:9055622
 A:Contents: annotation; bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: Kininogen is a cysteine proteinase inhibitor. takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is 1
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 xproline residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-References: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379, 390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <RCH>
 F:19-131/Domain: cystatin homology <CV1>
 F:142-253/Domain: cystatin homology <CT2>
 F:264-375/Domain: cystatin homology <CT3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:28-614, 83-94, 107-126, 142-145, 206-218, 229-248, 264-267, 328-340, 351-370/Disulfide bond
 F:48/Binding site: carboxydrate (Asn) (covalent) #status absent
 F:169, 205, 284/Binding site: carboxydrate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 579, 628/Binding site: carboxydrate (Thr) (covalent) #status
 F:577/Binding site: carboxydrate (Ser) (covalent) #status experimental

Query Match	Best Local Similarity	100.0%	Score 74;	DB 1;	Length 644;
Matches 12; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	GKKNGKHNGWKT 12			
DB	510	GKKNGKHNGWKT 521			

RESULT 2
 KGB0H2
 kininogen, HMW II precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A01282; A91923; A91943; A91938; B29559
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A:Title: A single gene for bovine high molecular weight and low molecular weight kin
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01282
 A:Molecule type: mRNA
 A:Residues: 1-619 <KIT>
 A:Cross-References: GB:V01492; GB:K01758; NID:9493; PIDN:CMA24736.1; PID:9494
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 376-391 <KAT>
 R:Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.

F:399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 68.9%; Score 51; DB 1; Length 621;
Best Local Similarity 66.7%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKKNKGKNGMKT 12
|| ||||| ||
DB 488 GKKNKGKHYDMKT 499

RESULT 4

T35064
probable integral membrane protein - streptomyces coelicolor

C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35064

R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999

A:Reference number: Z21567

A:Accession: T35064

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-275 <SEE>

A:Cross-references: EMBL:AL096884; PIDN:CA851427.1; GSPDB:GN00070; SCOEDB:SC4G6.04C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC4G6.04C

Query Match

Best Local Similarity 63.5%; Score 47; DB 2; Length 275;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 KKNKGNGW 10
|||||
DB 222 KKNKGNSW 229

RESULT 5

G71333
probable ribonuclease H (rnha) - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999

C:Accession: G71333

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; MCD
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770

A:Accession: G71333

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-169 <COLD>

A:Cross-references: GB:AE001215; GB:AE000520; NID:g3322631; PIDN:AAC65340.1; PID:g332263

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0353

C:Superfamily: ribonuclease H

Query Match 62.2%; Score 46; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 KKNKGKMT 12
|||||
DB 103 KKNKGKMT 109

RESULT 6
T49422
RAD57 related protein [imported] - Neurospora crassa

N:Alternate names: protein B17C10.30

C:Species: Neurospora crassa

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49422

R:Schulte, U.; Algn, V.; Hobeisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatu

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49422

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-516 <SCH>

A:Cross-references: EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.30

A:Experimental source: BAC clone B17C10; strain OR74A

C:Genetics:

A:Gene: NCSP:B17C10.30

A:Map position: 6

A:Introns: 31/3

Query Match 58.1%; Score 43; DB 2; Length 516;
Best Local Similarity 72.7%; Pred. No. 30;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKKNKGKNGMKT 11
|||||
DB 489 GKKNKGKNGMKT 499

RESULT 7

T11662
probable transcription factor 44K chain - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: T11662

R:Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z17305

A:Accession: T11662

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-374 <BAR>

A:Cross-references: EMBL:Z98533; NID:e1071719; PID:e334060

C:Genetics:

A:Map position: 1L

A:Note: SPAC6F12.11C

Query Match 56.8%; Score 42; DB 2; Length 374;
Best Local Similarity 72.7%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 KKNKGKNGMKT 12
|||||
DB 214 KKNKGKNGMKT 224

RESULT 8

T31081
ccas protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000

C:Accession: T31081

R:Hayashi, Y.; Ichinose, M.; Yuasa, H.; Tatematsu, M.; Ishibashi, M.

FEBS Lett. 406, 147-150, 1997

A:Title: Cca3, the mRNA level of which transiently decreases before initiation of DNA

A:Reference number: Z20971; MUID:97263491

A:Accession: T31081

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1009 <HAY>

A:Cross-references: EMBL:AB000216; NID:g2104557; PIDN:BA01969.1; PID:g2104558
C:Genetics:
A:Gene: cca3

Query Match 56.8%; Score 42; DB 2; Length 1009;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 GKKNGKHNGWKT 12
| | | | |
Db 59 GSNMNRHNSMDT 70

RESULT 9

ribonuclease H (EC 3.1.26.4) I NMA1817 [Imported] - Neisseria meningitidis (strain Z2491)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: F81807
R:Parthill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,
; Holroyd, S.; Jørgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; MUID:20222556
A:Accession: F81807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CA885042.1; PID:g738045
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: rnhA; NMA1817
A:Superfamily: ribonuclease H
C:Keywords: hydrolase

Query Match 55.4%; Score 41; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 9; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

OY 3 KNG-----KHNGWKT 12
| | | | |
Db 75 KNGMWHNGWKRNGWKT 92

RESULT 10

ribonuclease HI NMB1618 [Imported] - Neisseria meningitidis (strain MC58 serogroup B)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: H81061
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755
A:Accession: H81061
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <TEU>
A:Cross-references: GB:AE002512; GB:AE002098; NID:g7226866; PIDN:AA01970.1; PID:g722686
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1618
C:Superfamily: ribonuclease H

Query Match 55.4%; Score 41; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 9; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

OY 3 KNG-----KHNGWKT 12
| | | | |
Db 75 KNGMWHNGWKRNGWKT 92

RESULT 11

met101-1 protein - fruit fly (Drosophila hydei)
S34153
C:Species: Drosophila hydei
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34153
R:Neesen, J.; Heinlein, U.A.O.; Buenemann, H.
submitted to the EMBL Data Library, June 1993
A:Reference number: S34153
A:Accession: S34153
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-344 <NEE>
A:Cross-references: EMBL:X73480; NID:g313199; PID:g313200
C:Genetics:
A:Gene: FlyBase:Dhyd/met101
A:Cross-references: FlyBase:FBgn0011816
C:Superfamily: neurofilament triplet H protein

Query Match 55.4%; Score 41; DB 2; Length 344;
Best Local Similarity 72.7%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKKNGKHNGWKT 11
| | | | |
Db 304 GKKNGKKNMK 314

RESULT 12

S72275
nickase taxC - Escherichia coli plasmid R6K
N:Alternate names: DNA distortion protein 2
C:Species: Escherichia coli
C>Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 21-Jul-2000
C:Accession: S72275; S70937
R:Avila, P.; Nunez, B.; de la Cruz, F.
J. Mol. Biol. 261, 135-143, 1996
A>Title: Plasmid R6K contains two functional oris which can assemble simultaneously
A:Reference number: S72275; MUID:96346167
A:Accession: S72275
A:Molecule type: DNA
A:Residues: 1-385 <AVI>
A:Cross-references: EMBL:X95535; NID:g1524034; PIDN:CA064780.1; PID:g1524035
R:Plasmer, Y.; Shiomai, J.; Shaffer, A.
Mol. Microbiol. 19, 985-996, 1996
A>Title: Three novel plasmid R6K proteins act in concert to distort DNA within the pl
A:Reference number: S70936; MUID:96249692
A:Accession: S70937
A:Molecule type: DNA
A:Residues: 1-227, 'A', 229-255, 'K', 257-385 <FLA>
C:Genetics:
A:Gene: taxC
A:Map position: 20..3
A:Genome: plasmid R6K
C:Function:
A:Description: catalyzes a sequence-specific cleaving-joining reaction of single-str
C:Superfamily: Escherichia coli plasmid R6K nickase taxC
C:Keywords: DNA binding; plasmid DNA replication

Query Match 55.4%; Score 41; DB 2; Length 385;
Best Local Similarity 70.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 312 KNGIOGQWKT 321

RESULT 13

DB1290
probable capsule polysaccharide export system periplasmic protein Cj1444c [imported] - C
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C:Accession: DB1290
R:Parhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barré
Nature 403, 665-668, 2000
A>Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
A:Reference number: AB1250; MUID:20150912
A:Accession: DB1290
A:Status: preliminary.
A:Molecule type: DNA
A:Residues: 1-552 <PAR>
A:Cross-References: GB:AL139078; GB:AL111168; NID:96968723; PIDN:CAW3868.1; PID:9696887
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: kpsD; Cj1444c

Query Match 55.4%; Score 41; DB 2; Length 552;
Best Local Similarity 77.8%; Pred. No. 65;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 NGKHNGMKT 12
||:|||||
Db 329 NGKHNGMKT 337

RESULT 14

T05352
hypothetical protein F8B4.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05352
R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15409
A:Accession: T05352
A:Molecule type: DNA
A:Residues: 1-857 <BEV>
A:Cross-References: EMBL:AL034567
A:Experimental source: cultivar Columbia; BAC clone F8B4
C:Genetics:
A:Map position: 4
A:Introns: 26/3; 45/1; 74/3; 83/1; 122/2; 165/1; 270/2; 307/1; 731/2; 754/2
A>Note: F8B4.120
C:Superfamily: cyclophilin homology
F:6-162/Domain: cyclophilin homology <CYP>

Query Match 55.4%; Score 41; DB 2; Length 857;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GKKNGKH 7
|||||||
Db 171 GKKNGKH 177

RESULT 15

T27150
hypothetical protein Y54E2A.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T27150
R:Lloyd, C.
submitted to the EMBL Data Library, October 1998

A:Reference number: Z20319
A:Accession: T27150

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-198 <MTL>

A:Cross-References: EMBL:AL032646; PIDN:CAA21683.1; GSPDB:GN00019; CESP:Y54E2A.9

A:Experimental source: clone Y54E2A

C:Genetics:

A:Gene: CESP:Y54E2A.9

A:Map position: 1

A:Introns: 145/1

C:Superfamily: Caenorhabditis elegans hypothetical protein Y38H6C.19

Query Match 54.1%; Score 40; DB 2; Length 198;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKKNGKHNGM 10
||:|||||
Db 131 GKPNGKRNKM 140

Search completed: July 6, 2001, 09:18:00
Job time: 646 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:26:37 ; Search time 37.59 Seconds

(without alignments)
10.936 Million cell updates/sec

Title: US-09-437-912-4

Perfect score: 74

Sequence: 1 GKKNGKNGMKWT 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	51	68.9	619	1 KNH2_BOVIN	P01045 bos taurus
3	51	68.9	621	1 KNH1_BOVIN	P01044 bos taurus
4	46	62.2	169	1 KNH_TREPA	O83372 treponema p
5	42	56.8	374	1 YDVB_SCHPO	O14229 schizosacch
6	41	55.4	344	1 MST1_DROXY	O08695 dirosophila
7	39	52.7	304	1 R2_AQUAE	O06434 aquifex aeo
8	39	52.7	789	1 ACOX_YEAST	P35533 saccharomyc
9	39	52.7	874	1 POL1_HUMAN	P12666 homo sapien
10	38.5	52.0	491	1 NMT_CRYNE	P34809 cryptococcu
11	38	51.4	57	1 YMDP_ECOLI	P56614 escherichia
12	38	51.4	231	1 KNH_STRCO	O96776 streptomyc
13	38	51.4	578	1 GKR4_HUMAN	P33298 homo sapien
14	38	51.4	861	1 TOP1_BUCAL	P57371 buchnera ap
15	38	51.4	1534	1 MTDM_ARATH	P34881 arabidopsch
16	37.5	50.7	891	1 YB33_SCHPO	O14338 schizosacch
17	37	50.0	102	1 RK23_ODOST	P45559 odontella s
18	37	50.0	174	1 RN3_PORPU	P51298 porphyra pu
19	37	50.0	224	1 DCL_LYCES	O44663 lycopersico
20	37	50.0	263	1 VCP_VACVO	P10998 vaccinia vi
21	37	50.0	312	1 YAO9_SCHPO	O10108 schizosacch
22	37	50.0	357	1 ARHY_HUMAN	P54922 homo sapien
23	37	50.0	384	1 HEP1_PEDHE	O08819 pedobacter
24	37	50.0	408	1 YBX0_ARATH	O96934 arabidopsis
25	37	50.0	482	1 TPR_PORGI	P25806 porphyromon
26	37	50.0	515	1 PDI_ASPIG	O12730 aspergillus
27	37	50.0	552	1 HIS5_YEAST	P33734 saccharomyc
28	37	50.0	558	1 CABP_RAT	O65734 rattus norv
29	37	50.0	590	1 GKR5_BOVIN	P43249 bos taurus
30	37	50.0	590	1 GKR5_HUMAN	P34947 homo sapien
31	37	50.0	590	1 GKR5_HUMAN	O62833 rattus norv
32	37	50.0	686	1 MAS2_HUMAN	O00187 homo sapien
33	37	50.0	700	1 GLNA_BUTFI	O05650 butyrivibri

34	37	50.0	729	1 GLNA_BACFR	P15623 bacteroides
35	36.5	49.3	356	1 NTRB_RHOCA	P09431 rhodobacter
36	36	48.6	153	1 YGDI_HAEIN	P44831 haemophilus
37	36	48.6	185	1 PTH_RICPR	O92cv4 rickettsia
38	36	48.6	186	1 PTH_HELPJ	O92j33 helicobacte
39	36	48.6	186	1 PTH_HELPJ	P56077 helicobacte
40	36	48.6	450	1 VD10_BPT5	P11107 bacterioph
41	36	48.6	514	1 ABP2_BACOV	O59219 bacterioides
42	36	48.6	514	1 SAD1_SCHPO	O09825 schizosacch
43	36	48.6	518	1 IKAR_CHICK	O42410 gallus gall
44	36	48.6	642	1 YOR1_CAEEL	O09537 caenorhabdi
45	36	48.6	951	1 HEX_ADE05	P04133 human adeno

ALIGNMENTS

RESULT	ID	NAME	STANDARD	PRT	AA
1	KNG_HUMAN				
AC	P01042; P01043;				644 AA.
DT	21-JUL-1986	(Rel. 01, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ]				
GN	KNG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).			
RP	TISSUE=Liver;				
RC	MEDLINE=85234582; PubMed=2989293;				
RX	Takagaki Y., Kitamura N., Nakanishi S.;				
RA	"Cloning and sequence analysis of cDNAs for human high molecular weight and low molecular weight prekininogens. Primary structures of two human prekininogens.";				
RT	J. Biol. Chem. 260:8601-8609(1985).				
RL	[2]	GENE STRUCTURE.			
RN	MEDLINE=85234583; PubMed=2989294;				
RP	Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,				
RA	Nakanishi S.;				
RT	"Structural organization of the human kininogen gene and a model for its evolution.";				
RL	J. Biol. Chem. 260:8610-8617(1985).				
RN	[3]	SEQUENCE OF 1-401 FROM N.A.			
RP	MEDLINE=85122621; PubMed=6441591;				
RA	Ohkubo I., Kurauchi K., Takasawa T., Shiohara H., Sasaki M.;				
RT	"Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kininogen.";				
RL	Biochemistry 23:5691-5697(1984).				
RN	[4]	SEQUENCE OF 379-644.			
RP	MEDLINE=86030270; PubMed=4054110;				
RA	Lottspeich F., Kellermann J., Henschen A., Foertsch B.,				
RT	Mueller-Esterl W.;				
RL	"The amino acid sequence of the light chain of human high-molecular-mass kininogen.";				
RL	Eur. J. Biochem. 152:307-314(1985).				
RN	[5]	SEQUENCE OF 381-389.			
RP	MEDLINE=90255622; PubMed=4952632;				
RA	Pierce J.V.;				
RT	"Structural features of plasma kinins and kininogens.";				
RL	Fed. Proc. 27:52-57(1968).				
RN	[6]	DISULFIDE BONDS.			
RP	Sueyoshi T., Miyata T., Kato H., Iwanaga S.;				
RA	"Disulfide bonds in bovine HMW kininogens.";				

RL Selkagaku 56:808-808(1984).

CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)

CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY

CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO

CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-

CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE

CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF

CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE

CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND

CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A

CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR

CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF

CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS

CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,

CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)

CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-

CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD

CC CLOTTING.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: K02566; AAA35497.1; -

DR EMBL: M11437; AAB59550.1; -

DR EMBL: M11438; AAB59550.1; JOINED.

DR EMBL: M11521; AAB59550.1; JOINED.

DR EMBL: M11522; AAB59550.1; JOINED.

DR EMBL: M11523; AAB59550.1; JOINED.

DR EMBL: M11524; AAB59550.1; JOINED.

DR EMBL: M11525; AAB59550.1; JOINED.

DR EMBL: M11526; AAB59550.1; JOINED.

DR EMBL: M11527; AAB59550.1; JOINED.

DR EMBL: M11528; AAB59550.1; JOINED.

DR EMBL: M11529; AAB59551.1; -

DR EMBL: M11437; AAB59551.1; -

DR EMBL: M11438; AAB59551.1; JOINED.

DR EMBL: M11521; AAB59551.1; JOINED.

DR EMBL: M11522; AAB59551.1; JOINED.

DR EMBL: M11523; AAB59551.1; JOINED.

DR EMBL: M11524; AAB59551.1; JOINED.

DR EMBL: M11525; AAB59551.1; JOINED.

DR EMBL: M11526; AAB59551.1; JOINED.

DR EMBL: M11527; AAB59551.1; JOINED.

DR EMBL: M11528; AAB59551.1; JOINED.

DR PIR: A01279; KGHUHL.

DR PIR: A25276; A25276.

DR PIR: A01280; KGHUHL.

DR PIR: B25276; B25276.

DR PIR: S02482; S02482.

DR SWISS-2DPAGE: P01043; HUMAN.

DR MIM: 228960; -

DR InterPro: IPR000010; -

DR InterPro: IPR002395; -

DR Pfam: PF00031; cystatin.3.

DR PRINTS: PR00334; KININOGEN.

DR PROSITE: PS00287; CYSTATIN.2.

DR Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;

DR Bradykinin; Blood coagulation; Inflammatory response; Signal;

KW Alternative splicing.

FT SIGNAL 1 18

FT CHAIN 19 644 KININOGEN.

FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 389 BRADYKININ.

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD_RES	19	19	
FT	DISULFID	28	614	
FT	DISULFID	83	94	
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	
FT	CARBOHYD	169	169	
FT	CARBOHYD	205	205	
FT	CARBOHYD	294	294	
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPLIC	402	427	
FT	VARSPLIC	428	644	
FT	CONFLICT	593	593	
FT	SEQUENCE	644 AA;	71945 MW;	

Query Match 100.0%; Score 74; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKNGKNGMKWT 12
Db 510 GKNGKNGMKWT 521

RESULT 2
ID KNH2_BOVIN STANDARD; PRT; 619 AA.
AC P01045;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KININOGEN, HMW II PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS:
DE BRADYKININ].
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84014106; PubMed=6571699;
RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT "A single gene for bovine high molecular weight and low molecular
weight kininogens.";
RL Nature 305:545-549(1983).
RN [2]
RP SEQUENCE OF 19-376.
RX MEDLINE=87137530; PubMed=3546295;
RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,

VSPHPTSMAPADDERDSGKEGGHPR -> SHLRSCGYKGR
PKKGAEPASEREVS (IN ISOFORM LMW).
MISSING (IN ISOFORM LMW).
T -> I (IN REF. 1).
3132B4CBAF8FB7E CRC64;

RA Kato H., Nagasawa S., Suzuki T.;
 RT "Studies on the structure of bovine kininogen: cleavages of disulfide
 bonds and of methionyl bonds in kininogen-II.";
 RL J. Biochem. 67:313-323(1970).
 RN [4]

RP SEQUENCE OF 458-498.
 RX MEDLINE-75170265; PubMed-1169237;
 RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
 RT "Studies on the primary structure of bovine high-molecular-weight
 kininogen. Amino acid sequence of a fragment ('histidine-rich
 peptide') released by plasma kallikrein.";
 RL J. Biochem. 77:55-68(1975).

-1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOI PROTEASES; (2)
 HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
 TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
 PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
 PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
 A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
 MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
 NAURIRESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
 (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
 IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
 RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
 (4E) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
 ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
 ACTION).

-1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 -1- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
 PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
 SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
 TO RESIDUE 400.

-1- TISSUE SPECIFICITY: PLASMA.
 -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL: V01491; CAA24735.1; -
 DR PIR: A01281; KGB0H1.
 DR PIR: A29559; A29559.
 DR InterPro: IPR000010; -
 DR InterPro: IPR002395; -
 DR Pfam: PF00031; cystatin. 3.
 DR PRINTS: PR00334; KININOGEN.
 DR PROSITE: PS00287; CYSTATIN; 2.
 KW Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;
 KW Thiol protease inhibitor; Bradykinin; Blood coagulation;
 KW Inflammatory response; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 621
 FT CHAIN 19 378
 FT PEPTIDE 380 388
 FT CHAIN 389 621
 FT DOMAIN 19 135
 FT DOMAIN 136 257
 FT DOMAIN 258 378
 FT MOD_RES 19 19
 FT MOD_RES 87 87
 FT CARBOHYD 136 136
 FT CARBOHYD 136 136
 FT CARBOHYD 136 168
 FT CARBOHYD 158 197
 FT CARBOHYD 204 204
 FT CARBOHYD 27 591
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144

FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;

Query Match 68.9%; Score 51; DB 1; Length 621;
 Best Local Similarity 66.7%; Pred No. 0.51;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GKRNGKNGWKT 12
 Db 488 GKRNGKNGWKT 499

RESULT 4
 RNH_TREPA STANDARD; PRT; 169 AA.
 ID RNH_TREPA
 AC 083372;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE RIBONUCLEASE H (EC 3.1.26.4) (RNase H).
 GN RNH OR TP0353.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxId=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NICHOLS;
 RX MEDLINE-98332770; PubMed-9665876;
 RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterlidge T.,
 RA McDonald L., Atliach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 spirochete.";
 RL Science 281:375-388(1998).

-1- FUNCTION: THIS ENZYME IS AN ENDONUCLEASE THAT DEGRADES THE RNA OF
 CC RNA-DNA HYBRIDS SPECIFICALLY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
 CC MONESTER.
 CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY (BY SIMILARITY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE RNASE H FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL: AE001215; AAC65340.1; -
 DR TIGR: TP0353;
 DR InterPro: IPR002156; -
 DR Pfam: PF00075; rnaseh; 1.
 KW Hydrolase; Nuclease; Endonuclease; Magnesium.
 FT METAL 12 12
 FT METAL 63 63
 FT METAL 87 87
 FT METAL 151 151
 SO SEQUENCE 169 AA; 18184 MW; 1643110536328047 CRC64;

Query Match 62.2%; Score 46; DB 1; Length 169;

Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 KHNKMT 12
| | | | |
Db 103 KHNKMT 109

RESULT 5

YDVB_SCHPO STANDARD; PRT; 374 AA.
AC 014229;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHEICAL 43.4 KDA PROTEIN C6F12.11C IN CHROMOSOME I.
GN SPAC6F12.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; 298533; CAB11095.1; -
CC DR Hypothetical protein.
CC KW Hypothetical protein.
CC FT DOMAIN 66
CC FT SEQUENCE 374 AA; 43426 MW; 615C054F6AED0471 CRC64;
SQ

Query Match 56.8%; Score 42; DB 1; Length 374;
Best Local Similarity 72.7%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 KHNKHNKMT 12
| | | | |
Db 214 KHNKHNKMT 224

RESULT 6
MST1_DROHY STANDARD; PRT; 344 AA.
AC Q08695;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE AXONEME-ASSOCIATED PROTEIN MST101(1).
GN MST101(1).
OS Drosophila hydei (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephydroidea; Diptera; Ephydroidea; Diptera; Brachycera; Muscomorpha;
OX NCBI_TaxID=7224;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE-Testis;
RX MEDLINE=94200512; PubMed=8150205;
RA Neesen J., Buemann H., Heinlein U.A.;
RT "The Drosophila hydei gene Dmst101(1) encodes a testis-specific,
RT repetitive, axoneme-associated protein with differential abundance in
RT Y chromosomal deletion mutant flies.";

RL Dev. Biol. 162:414-425(1994).

CC -1- FUNCTION: POSSIBLE STRUCTURAL ROLE IN THE SPERM TAIL. IT IS
CC ASSOCIATED WITH AXONEMAL STRUCTURES.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: TESTIS. LOCATED IN SPERMATOCYTES AND
CC SPERMATID BUNDLES.
CC -1- DOMAIN: THE PREDOMINANT STRUCTURE IS ALPHA-HELICAL.
CC -1- POLYMORPHISM: THE NUMBER OF REPEATS VARIES BETWEEN STRAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X73480; CAA51875.1; -
CC DR PIR: S34153; S34153.
CC DR HSSP: P01032; 1C5A.
CC DR Flybase; FBgn0011816; Dhyd\mst101(1).
CC KW Sperm; Repeat; Multigene family.
CC FT DOMAIN 58
CC FT SEQUENCE 344 AA; 37793 MW; 24C65D2510387E2A CRC64;
SQ

Query Match 55.4%; Score 41; DB 1; Length 344;
Best Local Similarity 72.7%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKKNKHNKMT 11
| | | | |
Db 304 GKKNKHNKMT 314

RESULT 7
RI2_AQUAE STANDARD; PRT; 304 AA.
AC O6634;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L2.
GN RPLB OR AO_013.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: THIS PROTEIN IS A PRIMARY 23S RNA-BINDING PROTEIN. IT
CC HAS PEPTIDYLTRANSFERASE ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AE000669; AAC06392.1; -
CC DR InterPro; IPR002171; -
CC DR Pfam; PF00181; Ribosomal_L2; 1.

DR PROSITE; PS00467; RIBOSOMAL_L2; 1.
 KW Ribosomal protein; rRNA-binding.
 SQ SEQUENCE 304 AA; 34593 MW; 1052B90014979058 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 304;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 8; Conservative 1; Mismatches 3; Indels 4; Gaps 1;
 QY 1 GKNGKHNG---GWKT 12
 1: 111 1111
 DB 264 GTRGRKHSPSPGQWKT 279

RESULT 8
 ACOX_YEAST STANDARD; PRT; 789 AA.
 ID ACOX_YEAST
 AC P39533;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PUTATIVE ACETYLASE IN PRP21-UBP12 INTERGENIC REGION (EC 4.2.1.3).
 GN YUL200C OR J0337.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C;
 RX MEDLINE=95274326; PubMed=7754713;
 RA Purnelle B., Coster F., Goffeau A.;
 RT "The sequence of a 36 kb segment on the left arm of yeast chromosome
 RT X identifies 24 open reading frames including NC1, PRP21 (SEP91),
 RT CDC6, CRY2, the gene for S24, a homologue to the acetylase gene ACO1
 RT and two homologues to chromosome III genes."
 RT Yeast 10:1235-1249(1994).
 RL -1- CATALYTIC ACTIVITY: CITRATE = CIS-ACONITATE + H(2)O.
 CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X77688; CAA54757.1; -;
 DR EMBL; 249475; CAA89495.1; -;
 DR PIR; S46631; S46631.
 DR HSSP; P16276; GACN.
 DR SGD; S0003736; YUL200C.
 DR InterPro; IPR000573; -;
 DR InterPro; IPR001030; -;
 DR Pfam; PF00694; Aconitase.C; 1.
 DR Pfam; PF00330; Aconitase.1.
 DR PRINTS; PR00415; ACONITASE.
 DR PROSITE; PS00450; ACONITASE.1; 1.
 DR PROSITE; PS01244; ACONITASE.2; 1.
 KW Hypothetical protein; Lyase; Tricarboxylic acid cycle; Iron-sulfur;
 KW 4Fe-4S.
 FT METAL 385 385 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 448 448 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 451 451 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 SQ SEQUENCE 789 AA; 86563 MW; FABA4FE482D3F993 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 789;
 Best Local Similarity 63.6%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 KNGKHNGWKT 12

DB 230 KNGKHNGWKT 240

RESULT 9
 POLI_HUMAN STANDARD; PRT; 874 AA.
 ID POLI_HUMAN
 AC P10266;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE RETROVIRUS-RELATED POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE
 DE (EC 2.7.7.49); ENDONUCLEASE].
 GN POL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87036922; PubMed=3021993;
 RA Ono M., Yasunaga T., Miyata T., Ushikubo H.;
 RT "Nucleotide sequence of human endogenous retrovirus genome related to
 RT the mouse mammary tumor virus genome."
 RT J. Virol. 60:589-598(1986).
 RL PIR: D24483; GNMUR.
 DR HSSP; P03366; 1HMV.
 DR InterPro; IPR000477; -;
 DR InterPro; IPR001037; -;
 DR InterPro; IPR001584; -;
 DR InterPro; IPR002156; -;
 DR Pfam; PF02022; Integrase.Zn; 1.
 DR Pfam; PF00552; Integrase.1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00078; rvl; 1.
 KW Hydrolase; Transferase; RNA-directed DNA polymerase; Nuclease;
 KW Endonuclease; Polypeptide.
 FT CHAIN 36 250
 FT CHAIN 585 764 REVERSE TRANSCRIPTASE.
 FT CHAIN 585 764
 FT SEQUENCE 874 AA; 96936 MW; FD985989798018B6 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 874;
 Best Local Similarity 53.3%; Pred. No. 61;
 Matches 8; Conservative 0; Mismatches 3; Indels 4; Gaps 1;
 QY 1 GKNGKHNG---WK 11
 1111 1111
 DB 802 GKNGKHNGWKT 816

RESULT 10
 NMT_CRYNE STANDARD; PRT; 491 AA.
 ID NMT_CRYNE
 AC P34809;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLYCYLEPTIDE N-TETRADECANOTRANSFERASE (EC 2.3.1.97) (PEPTIDE
 DE N-MYRISTOYLTRANSFERASE) (MYRISTOYL-COA:PROTEIN N-MYRISTOYLTRANSFERASE)
 DE (NMT).
 OS Cryptococcus neoformans (Filobasidiella neoformans).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Tremellales;
 OC Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=5207;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L210425;
 RX MEDLINE=94132075; PubMed=8300631;
 RA Lodge J.K., Johnson R.L., Weinberg R.A., Gordon J.I.;
 RT "Comparison of Myristoyl-CoA:protein N-myristoyltransferases from
 RT three pathogenic fungi: Cryptococcus neoformans, Histoplasma

RT capsulatum, and Candida albicans.";
 RL J. Biol. Chem. 269:2996-3009(1994).
 CC -1- FUNCTION: ADDS MYRISTOYL GROUP TO N-TERMINAL GLYCINE RESIDUE OF
 CC CERTAIN CELLULAR PROTEINS.
 CC -1- CATALYTIC ACTIVITY: TETRADECANOYL-CoA + GLYCYL-PEPTIDE = CoA +
 CC N-TETRADECANOYLGLYCYL-PEPTIDE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE NMT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L25116; AAA17547.1; -;
 DR InterPro: IPR000903; -;
 DR Pfam: PF01233; NMT; 1.
 DR PROSITE: PS00975; NMT_1; 1.
 DR PROSITE: PS00976; NMT_2; 1.
 KM Transferase; Acyltransferase.
 FT MUTAGEN 487 487 G->D: REDUCED ACTIVITY.
 SQ SEQUENCE 491 AA: 54818 MW; BF84A3BC61349B9F CRC64;
 OY 1 GKKN-GKNGMKT 12
 Db 38 GTRNGEHEKMKWT 50
 OY 1 GKKN-GKNGMKT 12
 Db 38 GTRNGEHEKMKWT 50
 RESULT 11
 YMDF_ECOLI STANDARD: PRT; 57 AA.
 AC P56614;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHEICAL 5.9 KDA PROTEIN IN WRBA-PUTA INTERGENIC REGION.
 GN YMDF.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RA MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RL [2]
 RP IDENTIFICATION.
 RN Rudd K.E.;
 RA Unpublished observations (Aug-1998).
 CC -1- SIMILARITY: BELONGS TO THE CON-10 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL, AE000202; -; NOT_ANNOTATED_CDS.
 DR Ecogene; BGI4358; ymdf.
 KM Hypothetical protein.
 SQ SEQUENCE 57 AA: 5083 MW; D18133DAE91C6F5B CRC64;
 OY 1 GKKN-GKNGMKT 9
 Db 22 GKKGSGHSG 30
 OY 1 GKKN-GKNGMKT 9
 Db 22 GKKGSGHSG 30
 RESULT 12
 RNH_STRCO STANDARD: PRT; 231 AA.
 ID RNH_STRCO
 AC Q9X7R6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE RIBONUCLEASE H (EC 3.1.26.4) (RNASE H).
 GN RNHA OR SC5H1.08C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Oliver K., Harris D., James K.D., Parkhill J., Barrell B.G.,
 RA Rajandream M.A.;
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS ENZYME IS AN ENDONUCLEASE THAT DEGRADS THE RNA OF
 CC RNA-DNA HYBRIDS SPECIFICALLY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
 CC MONESTER.
 CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY (BY SIMILARITY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE RNASE H FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AL049863; CAB42933.1; ALT_INIT.
 DR InterPro: IPR002156; -;
 DR Pfam: PF00075; rnaesh; 1.
 KW Hydrolase; Nuclease; Endonuclease; Magnesium.
 FT METAL 10 10 MAGNESIUM (BY SIMILARITY).
 FT METAL 50 50 MAGNESIUM (BY SIMILARITY).
 FT METAL 72 72 MAGNESIUM (BY SIMILARITY).
 FT METAL 138 138 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 231 AA: 24297 MW; C650C69450C1899A CRC64;
 OY 6 KNGMKT 12
 Db 88 KNGMKT 94
 OY 6 KNGMKT 12
 Db 88 KNGMKT 94
 RESULT 13
 GKKA_HUMAN
 Query Match 51.4%; Score 38; DB 1; Length 231;
 Best Local Similarity 85.7%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ID GRK4_HUMAN STANDARD; PRT; 578 AA.
AC P32298; Q13293; Q13294; Q15313; Q15314; Q15315; Q15316;
AC 000641; 000642; 014453; 014725;
DT 01-OCT-1993 (Rel. 27, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE G PROTEIN-COUPLED RECEPTOR KINASE GRK4 (EC 2.7.1.-) (ITIL).
GN GPRK2L OR GPRK4 OR GRK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN MEDLINE FROM N.A. (GRK4-DELTA).
RX MEDLINE-93258311; PubMed-1338872;
RA Ambrose C., James M., Barnes G., Lin C., Bates G., Altherr M.,
RA Duyao M., Groot N., Church D., Wasmuth U.J., Lehrach H., Housman D.,
RA Buckler A.J., Gusella J.F., McDonald M.E.;
RT "A novel G protein-coupled receptor kinase gene cloned from 4p16.3";
RL Hum. Mol. Genet. 1:697-703(1992).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Testis;
RX MEDLINE-97248580; PubMed-9092566;
RA Salliese M., Mariglio S., Collodel G., Moretti E., Piomboni P.,
RA Baccetti B., de Biasi A.;
RT "G protein-coupled receptor kinase GRK4. Molecular analysis of the
RT four isoforms and ultrastructural localization in spermatozoa and
RT germinal cells";
RL J. Biol. Chem. 272:10188-10195(1997).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Testis;
RA Premont R.T., Lefkowitz R.J.;
RT "Characterization of the G protein-coupled receptor kinase GRK4.
RT Identification of four splice variants";
RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 3-83 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE-94183267; PubMed-8135832;
RA Salliese M., Lombardi M.S., de Biasi A.;
RT "Two isoforms of G protein-coupled receptor kinase 4 identified by
RT molecular cloning";
RL Biochem. Biophys. Res. Commun. 199:848-854(1994).
RN [5]
RP SEQUENCE OF 325-578 FROM N.A.
RA Gardner A.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: SPECIFICALLY PHOSPHORYLATES THE ACTIVATED FORMS OF G
CC PROTEIN-COUPLED RECEPTORS. GRK4-ALPHA CAN PHOSPHORYLATE RHODOPSIN
CC AND ITS ACTIVITY IS INHIBITED BY CALMODULIN. THE OTHER THREE
CC ISOFORMS DO NOT PHOSPHORYLATE RHODOPSIN AND DO NOT INTERACT WITH
CC CALMODULIN.
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; GRK4-ALPHA/GRK4D (SHOWN
CC HERE), GRK4-BETA/GRK4C, GRK4-GAMMA/GRK4A AND GRK4-DELTA/GRK4B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: TESTIS, AND IN A LOWER EXTENT IN OTHER TISSUES
CC INCLUDING BRAIN CORTEX AND STRIATUM.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC GPRK SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL; X97880; CAA66469.1; -
DR EMBL; X97881; CAA66470.1; -
DR EMBL; X33054; AAC50406.1; -
DR EMBL; X33055; AAC50407.1; -
DR EMBL; X33056; AAC50408.1; -
DR EMBL; X33168; AAC50409.1; -
DR EMBL; X33153; AAC50409.1; JOINED
DR EMBL; X33155; AAC50409.1; JOINED
DR EMBL; X33156; AAC50409.1; JOINED
DR EMBL; X33157; AAC50409.1; JOINED
DR EMBL; X33158; AAC50409.1; JOINED
DR EMBL; X33159; AAC50409.1; JOINED
DR EMBL; X33161; AAC50409.1; JOINED
DR EMBL; X33162; AAC50409.1; JOINED
DR EMBL; X33163; AAC50409.1; JOINED
DR EMBL; X33164; AAC50409.1; JOINED
DR EMBL; X33165; AAC50409.1; JOINED
DR EMBL; X33166; AAC50409.1; JOINED
DR EMBL; X33167; AAC50409.1; JOINED
DR EMBL; X33168; AAC50410.1; -
DR EMBL; X33153; AAC50410.1; JOINED
DR EMBL; X33154; AAC50410.1; JOINED
DR EMBL; X33155; AAC50410.1; JOINED
DR EMBL; X33156; AAC50410.1; JOINED
DR EMBL; X33157; AAC50410.1; JOINED
DR EMBL; X33158; AAC50410.1; JOINED
DR EMBL; X33159; AAC50410.1; JOINED
DR EMBL; X33161; AAC50410.1; JOINED
DR EMBL; X33162; AAC50410.1; JOINED
DR EMBL; X33163; AAC50410.1; JOINED
DR EMBL; X33164; AAC50410.1; JOINED
DR EMBL; X33165; AAC50410.1; JOINED
DR EMBL; X33166; AAC50410.1; JOINED
DR EMBL; X33167; AAC50410.1; JOINED
DR EMBL; X33168; AAC50411.1; -
DR EMBL; X33153; AAC50411.1; JOINED
DR EMBL; X33154; AAC50411.1; JOINED
DR EMBL; X33155; AAC50411.1; JOINED
DR EMBL; X33156; AAC50411.1; JOINED
DR EMBL; X33157; AAC50411.1; JOINED
DR EMBL; X33158; AAC50411.1; JOINED
DR EMBL; X33159; AAC50411.1; JOINED
DR EMBL; X33161; AAC50411.1; JOINED
DR EMBL; X33162; AAC50411.1; JOINED
DR EMBL; X33163; AAC50411.1; JOINED
DR EMBL; X33164; AAC50411.1; JOINED
DR EMBL; X33165; AAC50411.1; JOINED
DR EMBL; X33166; AAC50411.1; JOINED
DR EMBL; X33167; AAC50411.1; JOINED
DR EMBL; X33168; AAC50412.1; -
DR EMBL; X33153; AAC50412.1; JOINED
DR EMBL; X33154; AAC50412.1; JOINED
DR EMBL; X33155; AAC50412.1; JOINED
DR EMBL; X33156; AAC50412.1; JOINED
DR EMBL; X33157; AAC50412.1; JOINED
DR EMBL; X33158; AAC50412.1; JOINED
DR EMBL; X33159; AAC50412.1; JOINED
DR EMBL; X33161; AAC50412.1; JOINED
DR EMBL; X33162; AAC50412.1; JOINED
DR EMBL; X33163; AAC50412.1; JOINED
DR EMBL; X33164; AAC50412.1; JOINED
DR EMBL; X33165; AAC50412.1; JOINED
DR EMBL; X33166; AAC50412.1; JOINED
DR EMBL; X98118; CAA66802.1; -
DR EMBL; X75897; CAA53506.1; -
DR EMBL; X68192; CAA92341.1; -
DR HSP; 063450; 1A06.
DR MIM; 137026; -
DR InterPro; IPR000239; -
DR InterPro; IPR000342; -
DR InterPro; IPR000719; -


```
DR InterPro: IPR002290; -
DR Pfam: PF00615; RGS; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00717; GPCRINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00132; RGS; 1.
DR Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing; Polymorphism.
FT DOMAIN 1 154
FT N-TERMINAL.
FT DOMAIN 52 172
FT RGS.
FT DOMAIN 187 449
FT PROTEIN KINASE.
FT DOMAIN 450 578
FT C-TERMINAL.
FT NP_BIND 193 201
FT ATP (BY SIMILARITY).
FT BINDING 216 216
FT ATP (BY SIMILARITY).
FT ACT_SITE 312 312
FT MISSING (IN ISOFORM GRK4-BETA AND ISOFORM
FT VARSPLIC 18 49
FT GRK4-DELTA).
FT VARSPLIC 516 561
FT MISSING (IN ISOFORM GRK4-GAMMA AND
FT VARSPLIC 247 247
FT ISOFORM GRK4-DELTA).
FT V->I.
FT VARIANT 247 247
FT /FTID=VAR_007806.
FT CONFLICT 562 562
FT G->D (IN REF. 1, CAA66468 AND
FT CAA66802).
SQ SEQUENCE 578 AA; 66555 MW; 2CF2075EC336E4A1 CRC64;
```

```
Query Match 51.4%; Score 38; DB 1; Length 578;
Best Local Similarity 54.5%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 GKNGKNGMK 11
111:11 11
DB 21 GKNGKNGMK 31
```

RESULT 14

```
TOP1_BUCAT STANDARD; PRT; 861 AA.
AC P57371;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA TOPOISOMERASE I (EC 5.99.1.2) (OMEGA-PROTEIN) (RELAXING ENZYME)
DE (UNWINDING ENZYME) (SWIHELASE).
GN TOPA OR BU284.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TORO 1998;
RA MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS. ";
RL Nature 407:81-86(2000).
CC -1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.
CC -1- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED
CC DNA, FOLLOWED BY PASSAGE AND REJOINING.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
DR EMBL: AP001118; BAB12994.1; -
DR PROSITE: PS00396; TOPOISOMERASE_I_PROK; 1.
KW Isomerase; Topoisomerase; DNA-binding; Zinc-finger.
FT ACT_SITE 318 318
FT DNA_CLEAVAGE (BY SIMILARITY).
FT ZN_FING 596 628
FT C4-TYPE (POTENTIAL).
FT ZN_FING 658 685
FT C4-TYPE (POTENTIAL).
FT ZN_FING 707 732
FT C4-TYPE (POTENTIAL).
SQ SEQUENCE 861 AA; 99739 MW; 903BBF2EA5265599 CRC64;
```

```
Query Match 51.4%; Score 38; DB 1; Length 861;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 KNGKNGMW 10
11:11 11
DB 839 KNGKNGMW 847
```

RESULT 15

```
MTDM_ARATH STANDARD; PRT; 1534 AA.
AC P34861;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA (CYTOSINE-5)-METHYLTRANSFERASE ATH1 (EC 2.1.1.37) (DNA
DE METHYLTRANSFERASE ATH1) (DNA METHASE ATH1) (M.ATH1).
GN ATH1M.
OS Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA MEDLINE=93281384; PubMed=8389441;
RA Finnegan E.J., Dennis E.S.;
RT Isolation and identification by sequence homology of a putative
RT cytosine methyltransferase from Arabidopsis thaliana. ";
RL Nucleic Acids Res. 21:2383-2388(1993).
CC -1- FUNCTION: METHYLATES CG RESIDUES.
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA = S-ADENOSYL-L-
CC HOMOCYSTEINE + DNA CONTAINING 5-METHYLCYTOSINE.
CC -1- SIMILARITY: LOW, TO OTHER EUKARYOTIC DNA METHASE.
CC -1- SIMILARITY: SOME TO BACTERIAL RESTRICTION SYSTEMS
CC METHYLTRANSFERASES.
CC -1- SIMILARITY: CONTAINS 2 BAH DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
EMBL: L10692; AAA32829.1; -
REBASE: 2839; M.Ath1.
DR InterPro: IPR001025; -
DR InterPro: IPR001525; -
DR Pfam: PF01426; BAH; 2.
DR Pfam: PF01445; DNA_methylase; 3.
DR PRINTS: PR00105; C5METHFRASE.
DR PROSITE: PS00094; C5_MTASE_1; 1.
DR PROSITE: PS00095; C5_MTASE_2; 1.
```

KW Transferase; Methyltransferase; DNA-binding.
 FT ACT_SITE 1198 BY SIMILARITY.
 SQ SEQUENCE 1534 AA; 172430 MW; 23FC944AA7074C5A CRC64;

Query Match 51.4%; Score 38; DB 1; Length 1534;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KHNGWK 11
 :|||||
 Db 1434 KHNGWK 1439

Search completed: July 6, 2001, 09:26:38
 Job time: 969 sec

DB 222 KNGKHNW 229

RESULT 2

ID 068965 PRELIMINARY; PRT; 3898 AA.

AC 068965;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE HOG CHOLERA VIRUS POLYPROTEIN.

OS Classical swine fever virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Pestivirus.

OX NCBI_TaxID=11096;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-GPE-;

RX MEDLINE-95390794; PubMed-7661692;

RA Ishikawa K., Nagai H., Katayama K., Tsutsui M., Tanabayashi K.,

RA Takeuchi K., Hishiyama M., Saitoh A., Takagi M., Gotoh K.,

RA Muramatsu M., Yamada A.;

RT "Comparison of the entire nucleotide and deduced amino acid sequences

RT of the attenuated hog cholera vaccine strain GPE- and the wild-type

RT parental strain ALD-";

RT Arch. Virol. 140:1385-1391(1995).

RC -1 SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.

DR EMBL: D49533; BAA08477.1; -.

DR InterPro: IPR000280; -.

DR InterPro: IPR001005; -.

DR InterPro: IPR001410; -.

DR InterPro: IPR001568; -.

DR InterPro: IPR001650; -.

DR Pfam: PF00271; Helicase_C.1.

DR PRINTS: PR00729; CDVENDOPTASE.

DR PROSITE: PS00531; RNASE_T2_2; UNKNOWN_1.

DR SMART: SM00490; HELIC; 1.

DR ATP-binding; Helicase; Polyprotein.

SQ SEQUENCE 3898 AA; 438268 MW; D167BF5E48B11747 CRC64;

Query Match 60.8%; Score 45; DB 14; Length 3898;

Best Local Similarity 70.0%; Pred. No. 97;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GKKNGKHNGM 10

DB 3270 GSSLGKHNGM 3279

RESULT 3

ID 09P6E6 PRELIMINARY; PRT; 516 AA.

AC 09P6E6;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE RAD57 RELATED PROTEIN.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariales; Sordariaceae; Neurospora.

OX NCBI_TaxID=5141;

RN [1]

RP SEQUENCE FROM N.A.

RA Schulte U., Algen V., Hoheisel J., Brandt P., Farlmann B., Holland R.,

RA Nyakatura G., Mewes H.W., Mannhaupt G.;

RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA German Neurospora genome project;

RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL355926; CAB91223.1; -.

DR InterPro: IPR001553; -.

DR InterPro: IPR003593; -.

DR SMART: SM00382; AAA; 1.

SQ SEQUENCE 516 AA; 55456 MW; 8A1B9CA95D6AA295 CRC64;

Query Match 58.1%; Score 43; DB 3; Length 516;

Best Local Similarity 72.7%; Pred. No. 26;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GKKNGKHNGM 11

DB 489 GKKNGKEGKM 499

RESULT 4

ID 09UT73 PRELIMINARY; PRT; 171 AA.

AC 09UT73;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE MRNA, PARTIAL CDS (FRAGMENT).

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RA Kawamukai M.;

RT "S. pombe unknown protein.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB032716; BAA04655.1; -.

FT NON_TER 1

SQ SEQUENCE 171 AA; 19913 MW; CB35800E91F17E0D CRC64;

Query Match 56.8%; Score 42; DB 3; Length 171;

Best Local Similarity 72.7%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKKNGKHNGM 12

DB 11 KKKNGKHNGM 21

RESULT 5

ID 09DRT6 PRELIMINARY; PRT; 205 AA.

AC 09DRT6;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE NEF PROTEIN.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-LTS 21E;

RA Ashton L., Rhodes D., Solomon A., Deacon N., Satchell C., Carr A.,

RA Cooper D., Bitt R., Stewart G., Kaldor J.;

RT "Viral diversity in the nef/LTR region of the HIV-1 genome:

RT associations with long-term nonprogression.";

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF219695; AAC44172.1; -.

SQ SEQUENCE 205 AA; 23370 MW; 63BB87C97291AB2D CRC64;

Query Match 56.8%; Score 42; DB 14; Length 205;

SQ SEQUENCE 739 AA; 84293 MW; F6B92B3B8DC74216 CRC64;

RA Haft D.H., Salzburg S.L., White O., Fleischmann R.I.

RA Eissen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Hatt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clittone H., Clark E.B.,

RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathavan J.,
 RA Gill J., Scariato V., Masignani V., Pizsa M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappoli R., Venter J.C.;
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 RT MC58";
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002512; AAF41970.1; -;
 DR TIGR: NMB1618; -;
 DR InterPro: IPR002156; -;
 DR Pfam: PF00075; rnaaseH; 1.
 SQ SEQUENCE 145 AA; 16251 MW; 3E2EA6BED17D5E91 CRC64;

Query Match 55.4%; Score 41; DB 2; Length 145;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 9; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

OY 3 KNG-----KNGWKT 12
 ||| | |||||
 DB 75 KNGMENWIGHGKNGWKT 92

RESULT 10
 O9JTD9 PRELIMINARY; PRT; 145 AA.
 AC O9JTD9;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE RIBONUCLEASE HI (EC 3.1.26.4).
 GN KNAH OR NMA1817.
 OS *Neisseria meningitidis* (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
 RC MEDLINE=20222556; PUBMED=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
 RA Jagsels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajadaram M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of *Neisseria*
 RT meningitidis 22491";
 RL Nature 404:502-506(2000).
 DR EMBL: AL162757; CAB85042.1; -;
 DR InterPro: IPR002156; -;
 DR Pfam: PF00075; rnaaseH; 1.
 KW Hydrolase.
 SQ SEQUENCE 145 AA; 16235 MW; 3E2EA6A1317D4171 CRC64;

Query Match 55.4%; Score 41; DB 2; Length 145;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 9; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

OY 3 KNG-----KNGWKT 12
 ||| | |||||
 DB 75 KNGMENWIGHGKNGWKT 92

RESULT 11
 O9DRT9 PRELIMINARY; PRT; 208 AA.
 AC O9DRT9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE NEF PROTEIN.
 GN NEF
 OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LTS 21B;
 RA Ashton L., Rhodes D., Solomon A., Deacon N., Satchell C., Carr A.,
 RA Cooper D., Bitt R., Stewart G., Kaldor J.;
 RT "Viral diversity in the nef/LTR region of the HIV-1 genome:
 RT associations with long-term nonprogression";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF219692; AAG44169.1; -;
 SQ SEQUENCE 208 AA; 23738 MW; 9D55ACEB3832763FB CRC64;

Query Match 55.4%; Score 41; DB 14; Length 208;
 Best Local Similarity 54.5%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKNGKNGWKT 11
 | | | | |
 DB 2 GKNGKNGWKT 12

RESULT 12
 O9DRT8 PRELIMINARY; PRT; 208 AA.
 AC O9DRT8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE NEF PROTEIN.
 GN NEF
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LTS 21C;
 RA Ashton L., Rhodes D., Solomon A., Deacon N., Satchell C., Carr A.,
 RA Cooper D., Bitt R., Stewart G., Kaldor J.;
 RT "Viral diversity in the nef/LTR region of the HIV-1 genome:
 RT associations with long-term nonprogression";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF219693; AAG44170.1; -;
 SQ SEQUENCE 208 AA; 23738 MW; 5B9560E7D4AB85D7 CRC64;

Query Match 55.4%; Score 41; DB 14; Length 208;
 Best Local Similarity 54.5%; Pred. No. 22;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKNGKNGWKT 11
 | | | | |
 DB 2 GKNGKNGWKT 12

RESULT 13
 O9L2L5 PRELIMINARY; PRT; 359 AA.
 AC O9L2L5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PHO-LIKE PROTEIN.
 GN SCC117.05.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Oliver K., Harris D.;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kiser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL136534; CAB66426.1; -.
 SQ SEQUENCE 359 AA; 39146 MW; EAB26376BF0E45C7 CRC64;

Query Match 55.4%; Score 41; DB 2; Length 359;
 Best Local Similarity 58.3%; Pred. No. 39;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GKNGKNGMKT 12
 | : | | | : |
 DB 338 GTENGCHNGRT 349

RESULT 14
 Q9FHA0 PRELIMINARY; PRT; 365 AA.
 AC Q9FHA0;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE SIMILARITY TO MAPK-LIKE PROTEIN KINASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:31-63(2000).
 DR EMBL: AB020742; BAB10947.1; -.
 KW Kinase.
 SQ SEQUENCE 365 AA; 42010 MW; 0355B09EDE7F9DCC CRC64;

Query Match 55.4%; Score 41; DB 10; Length 365;
 Best Local Similarity 72.7%; Pred. No. 39;
 Matches 8; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

OY 3 KNGKH-NGWK 11
 | | | | |
 DB 288 KNSKHGNGWK 298

RESULT 15
 ID P71232 PRELIMINARY; PRT; 385 AA.
 AC P71232;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE NICKASE.

GN TAXC.
 OS Escherichia coli.
 OG Plasmid R6K.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96346167; PubMed=8757282;
 RA Avila P., Nunez B., De la Cruz F.;
 RT "Plasmid R6K contains two functional origins which can assemble
 RT simultaneously in relaxosomes in vivo.";
 RL J. Mol. Biol. 261:135-143(1996).
 DR EMBL: X95535; CAA64780.1; -.
 KW Plasmid.
 SQ SEQUENCE 385 AA; 44116 MW; 20AB3E8342C8C17 CRC64;

Query Match 55.4%; Score 41; DB 2; Length 385;
 Best Local Similarity 70.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 KNGKHNGMKT 12
 | | | | |
 DB 312 KNGIOGNGWK 321

Search completed: July 6, 2001, 09:25:54
 Job time: 990 sec


```
; Patent No. 6218147
; GENERAL INFORMATION:
; APPLICANT: LINGWOOD, Clifford A.
; TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/456,287
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/686,528
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rea, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 032609-001
; INFORMATION FOR SEQ. ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-456-287-3

Query Match          34.1%; Score 57; DB 4; Length 313;
Best Local Similarity 37.0%; Pred. No. 1.4;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 HGHEDQHGLGHGKFKLDDLEHGQGH 27
    | : | | | | | | | | | | : |
DB 101 HDHDKHKKHDKHDKHDDHDKHKKH 127

RESULT 3
US-08-686-528A-2
; Sequence 2, Application US/08686528A
; Patent No. 6054134
; GENERAL INFORMATION:
; APPLICANT: LINGWOOD, Clifford A.
; TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,528A
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Rea, Teresa Stanek
```

```
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 032609-001
; INFORMATION FOR SEQ. ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-686-528A-2

Query Match          34.1%; Score 57; DB 3; Length 337;
Best Local Similarity 37.0%; Pred. No. 1.5;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 HGHEDQHGLGHGKFKLDDLEHGQGH 27
    | : | | | | | | | | | | : |
DB 125 HDHDKHKKHDKHDKHDDHDKHKKH 151

RESULT 4
US-09-456-287-2
; Sequence 2, Application US/09456287
; Patent No. 6218147
; GENERAL INFORMATION:
; APPLICANT: LINGWOOD, Clifford A.
; TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/456,287
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/686,528
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rea, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 032609-001
; INFORMATION FOR SEQ. ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-456-287-2

Query Match          34.1%; Score 57; DB 4; Length 337;
Best Local Similarity 37.0%; Pred. No. 1.5;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 HGHEDQHGLGHGKFKLDDLEHGQGH 27
    | : | | | | | | | | | | : |
DB 125 HDHDKHKKHDKHDKHDDHDKHKKH 151

RESULT 5
US-07-945-283-2
```

```
; Sequence 2, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The Ep0 and LfT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4128
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1958 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-945-283-2

Query Match          31.1%; Score 52; DB 1; Length 1958;
Best Local Similarity 46.7%; Pred. No. 52;
Matches 14; Conservative 2; Mismatches 6; Indels 8; Gaps 3;

QY      1 HGHEOHGLGHG--HKFKLD--DDLEHOG 25
      11 :||||| 11 ||:| 1
Db      135 HG---EHGLHALAVHVLDDMDVYHVG 161

RESULT
US-08-446-345-36.
; Sequence 36, Application US/08446345
; Patent No. 5831009
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; TITLE OF INVENTION: NOVEL PROTEIN PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASES PTP-DI
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: N.Y.
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,345
```

```
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,440
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 7683-054
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNTE
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1174 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-446-345-36

Query Match          29.9%; Score 50; DB 2; Length 1174;
Best Local Similarity 30.8%; Pred. No. 55;
Matches 12; Conservative 4; Mismatches 9; Indels 14; Gaps 1;

QY      2 GHEOHGLGHGKFKL-----DDLEHOG 26
      1 :|:|:| 1
Db      686 GPEAEGLRYGHKKSLSDATMLHSSEDEDEFEESG 724

RESULT
US-08-937-931-2
; Sequence 2, Application US/08937931
; Patent No. 5935792
; GENERAL INFORMATION:
; APPLICANT: Rubin, Gerald M.
; APPLICANT: Pan, Duojia
; APPLICANT: Rooke, Jenny
; APPLICANT: Yavari, Reza
; APPLICANT: Xu, Tian
; TITLE OF INVENTION: KUZ: A No. 5935792el Family of Metalloproteases
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,931
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B97-081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1239 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

MOLECULE TYPE: peptide
US-08-937-931-2

Query Match 29.0%; Score 48.5; DB 2; Length 1239;
Best Local Similarity 37.0%; Pred. No. 94;
Matches 10; Conservative 2; Mismatches 10; Indels 5; Gaps 1;

OY 1 HGEHQHGLGHGKFKLDDLEHOGH 27
Db 1004 HGGSRSHHOHHPHDWD-----RHOGGH 1025

RESULT 8

US-09-285-502-2
Sequence 2, Application US/09285502
Patent No. 6190876
GENERAL INFORMATION:
APPLICANT: Rubin, Gerald M.
APPLICANT: Pan, Duojia
APPLICANT: Rooke, Jenny
APPLICANT: Yavari, Reza
APPLICANT: Xu, Tian
TITLE OF INVENTION: KUZ. A No. 6190876el Family of Metalloproteases
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/285,502
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,931
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-081
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-285-502-2

Query Match 29.0%; Score 48.5; DB 4; Length 1239;
Best Local Similarity 37.0%; Pred. No. 94;
Matches 10; Conservative 2; Mismatches 10; Indels 5; Gaps 1;

OY 1 HGEHQHGLGHGKFKLDDLEHOGH 27
Db 1004 HGGSRSHHOHHPHDWD-----RHOGGH 1025

RESULT 9
US-08-255-457-1
Sequence 1, Application US/08255457

Patent No. 5780040
GENERAL INFORMATION:
APPLICANT: Plant, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.
APPLICANT: Wright, Andrew
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/255,457
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-255-457-1

Query Match 28.7%; Score 48; DB 1; Length 60;
Best Local Similarity 40.0%; Pred. No. 3.8;
Matches 10; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

OY 3 HGEHQHGLGHGKFKLDDLEHOGH 27
Db 4 HGEHQHGLGHGKFKLDDLEHOGH 28

RESULT 10
US-09-115-032-1
Sequence 1, Application US/09115032
Patent No. 5972348
GENERAL INFORMATION:
APPLICANT: Plant, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.
APPLICANT: Wright, Andrew
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

MOLECULE TYPE: protein

US-08-584-031-8

REFERENCE/DOCKET NUMBER: P0978P1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5416
TELEFAX: 415/952-9881
TELETYPE: 910/271-7168
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-780-496-8

Query Match 27.8%; Score 46.5; DB 3; Length 24;
Best Local Similarity 39.1%; Pred. No. 2.2;
Matches 9; Conservative 2; Mismatches 7; Indels 5; Gaps 1;

QY 1 HGHEQHGIGHGKFKLDDLEH 23
DB 6 HHHHHHSSGH-----IDDDKH 23

RESULT 14
US-08-137-614A-26

Sequence 26, Application US/08137614A
Patent No. 5487976
GENERAL INFORMATION:
APPLICANT: Soderlund, David M.
APPLICANT: Kolipole, Douglas C.
APPLICANT: Henderson, Joseph E.
TITLE OF INVENTION: Gene Encoding An Insect
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
SMREER: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,614A
FILING DATE: 15-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1636
TELEFAX: (716)263-1600
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-137-614A-26

Query Match 27.8%; Score 46.5; DB 1; Length 617;
Best Local Similarity 38.5%; Pred. No. 81;
Matches 10; Conservative 1; Mismatches 8; Indels 7; Gaps 1;

QY 2 GHEQHGIGHGKFKLDDLEHGGH 27
DB 432 GMPREHGHGHH-----AHSHGH 450

RESULT 15
US-08-072-064-1
Sequence 1, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FRENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,064
FILING DATE: 19930602
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 770,881
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OPND-00574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster
POSITION IN GENOME:
CHROMOSOME/SEGMENT: III; polytene subregion 66F
MAP POSITION: approximately map unit 26
US-08-072-064-1

Query Match 27.8%; Score 46.5; DB 3; Length 637;
Best Local Similarity 38.5%; Pred. No. 84;
Matches 10; Conservative 1; Mismatches 8; Indels 7; Gaps 1;

QY 2 GHEQHGIGHGKFKLDDLEHGGH 27
DB 453 GMPREHGHGHH-----AHSHGH 471

Search completed: July 6, 2001, 09:10:22
Job time: 188 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:18:00 ; Search time 73.59 Seconds

(without alignments)
28.983 Million cell updates/sec

Title: US-09-437-912-5

Perfect score: 167

Sequence: 1 HGHQDQHGHLGHKFKLDDLEHGGHCV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167	100.0	644	1 KGHUHI	kininogen, HMW pre
2	98.5	59.0	264	2 C25486	K-kininogen, HMW p
3	98.5	59.0	639	2 A23486	kininogen, HMW I p
4	91.5	54.8	290	2 C27115	K-kininogen, LMW p
5	91.5	54.8	315	2 A27115	major acute phase
6	85	50.9	619	1 KGBOH2	kininogen, HMW II
7	85	50.9	621	1 KGBOH1	kininogen, HMW I p
8	65	38.9	2038	2 A43742	female sterile hom
9	64	38.3	776	2 T02702	hypothetical prote
10	63.5	38.0	436	2 T02702	MHC H-2K/L-w5-link
11	62.5	37.4	535	2 S66148	gene pipsqueak pro
12	62.5	37.4	1085	2 S66149	gene pipsqueak pro
13	62	37.1	314	2 T35241	hypothetical prote
14	61.5	36.8	199	2 T48099	hypothetical prote
15	61	36.5	735	2 T45059	hypothetical prote
16	60.5	36.2	398	2 T02681	probable zinc tran
17	60.5	35.9	515	2 T23089	hypothetical prote
18	60	35.9	191	2 D66701	hypothetical prote
19	59.5	35.6	201	2 H82055	unknown protein, 9
20	59	35.3	110	2 T07618	peptidyl prolyl ci
21	59	34.7	457	2 S39079	cold stress protei
22	58	34.7	378	2 T49164	put C-8 protein -
23	58	34.7	549	2 T15506	zinc transporter-1
24	57	34.1	213	2 T36345	hypothetical prote
25	57	34.1	335	2 D38532	hypothetical prote
26	57	34.1	337	2 D64045	hypB protein - rho
27	57	34.1	337	2 D64045	adhesin homolog HI
28	56.5	33.8	254	2 A31488	hypothetical prote
29	56.5	33.8	313	2 A28444	filaggrin - mouse

30	56	33.5	439	2 S58327	cobalt accumulatio
31	55.5	33.2	102	2 T30119	hypothetical prote
32	55.5	33.2	242	2 H82061	hypothetical prote
33	55.5	33.2	250	2 B35026	filaggrin B - mous
34	55.5	33.2	255	2 A35026	filaggrin A - mous
35	55	32.9	189	2 C81428	peptidyl-prolyl ci
36	55	32.9	283	2 C85838	hypothetical prote
37	54.5	32.6	410	2 T26757	hypothetical prote
38	54	32.3	389	2 B96635	hypothetical prote
39	53.5	32.0	258	2 A46820	homeobox protein (
40	53	31.7	18	2 B32473	histidine-rich pro
41	53	31.7	85	2 A45969	hemolymph antifung
42	53	31.7	245	2 T23844	hypothetical prote
43	53	31.7	409	2 E83992	ATP/GTP-binding pr
44	53	31.7	529	2 T08684	hypothetical prote
45	53	31.7	1061	2 S37667	trac-1 protein - E

ALIGNMENTS

RESULT 1
KGHUI
kininogen, HMW precursor [validated] - human
N:Alternate names: alpha-2-thiol proteinase inhibitor; proprokinogen; prokininogen
M:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C:Accession: A01279; A25276; S32422; A24871; A27899; A31905; A34030;
R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A:Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:85122621
A:Accession: A01279
A:Molecule type: mRNA
A:Residues: 1-389 <OHK>
A:Cross-references: GB:K02566; NID:9177889
R:Takagaki, Y.; Kitamura, N.; Nakamishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A:Title: Cloning, expression and sequence analysis of cDNAs for human high molecular weight and 1
A:Reference number: A92544; MUID:85234582
A:Accession: A25276
A:Molecule type: mRNA
A:Residues: 1-592, 'I', 594-644 <TAK>
A:Cross-references: GB:M1437; NID:9186751; PID:NAB59550.1; PID:9386852
R:Auerwald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A:Title: Cloning, expression and characterization of human kininogen domain 3.
A:Reference number: S32422; MUID:9323854
A:Accession: S32422
A:Molecule type: mRNA
A:Residues: ANSM, 253-377 <AUE>
A:Note: differences are due to known cloning artifacts
R:Lottspesch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A:Title: The amino acid sequence of the light chain of human high-molecular-mass kini
A:Reference number: A91153; MUID:86030270
A:Accession: A91153
A:Molecule type: protein
A:Note: the bradykinin sequence preceding the light chain sequence was not determined
R:Kellermann, J.; Lottspesch, F.; Henschen, A.; Muller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A:Title: Completion of the primary structure of human high-molecular-mass kininogen.
A:Reference number: A24871; MUID:86108361
A:Accession: A24871
A:Molecule type: protein
A:Residues: 'I', 20-380 <KEII>
R:Kellermann, J.; Lottspesch, F.; Henschen, A.; Muller-Esterl, W.
In Kinins IV, Greenbaum, U.M., and Margolis, H.S., ed., pp.85-89, Plenum Press, New
A:Title: Amino acid sequence of the light chain of human high molecular mass kininogen
A:Reference number: A27899
A:Accession: A27899

A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <REL2>
 R:Minotoli, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88209021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
 A:Reference number: A31905; MUID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAE>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plas
 A:Reference number: A34030; MUID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Renneric, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
 A:Reference number: S02482; MUID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19;189-192;310-314;381-389 <LENI>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
 A:Reference number: A61495; MUID:88211869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT3>
 R:Renneric, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
 FEBS Lett. 280, 211-215, 1991
 A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
 A:Reference number: S14303; MUID:91192133
 A:Accession: S14447
 A:Molecule type: protein
 A:Residues: 264-359, 'N', 361-375 <LEN2>
 R:Little, S.S.; Johnson, D.A.
 Biochem. J. 307, 341-346, 1995
 A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
 A:Reference number: S55239; MUID:95251593
 A:Accession: S55239
 A:Molecule type: protein
 A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
 R:Straczek, J.; Macchi, F.; Le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabec, P.; Bellevil
 FEBS Lett. 373, 207-211, 1995
 A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
 A:Reference number: S68059; MUID:96033974
 A:Accession: S68059
 A:Molecule type: protein
 A:Residues: 431-434 <STR>
 R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evoluti
 A:Reference number: A92545; MUID:85234583
 A:Contents: annotation; gene organization

R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:90255622
 A:Contents: annotation; bradykinin
 C:Comment: The HMW kininogen precursor and the LMW form are produced from the same ge
 C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is 1
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 C:Comment: xproline residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status experimental <Sig>
 F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
 F:19-379, 390-644/Product: HMW kininogen II #status experimental <MAT2>
 F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
 F:19-131/Domain: cystatin homology <CY1>
 F:142-253/Domain: cystatin homology <CY2>
 F:264-375/Domain: cystatin homology <CY3>
 F:280-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
 F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
 F:431-434/Product: low molecular weight growth promoting factor #status experimental
 F:19-614/Binding site: pyrrolidone carboxylic acid (Gln) (in mature form) #status expert
 F:28-614, 83-94, 107-126, 142-145, 206-218, 229-246, 264-267, 328-340, 351-370/Disulfide Bond
 F:48/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:169, 205, 294/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401, 533, 542, 546, 557, 571, 593, 628/Binding site: carbohydrate (Thr) (covalent) #status
 F:577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 167; DB 1; Length 644;
 Best Local Similarity 100.0%; Pred. No. 5.8e-15;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGHEQOHGIGHGKFKRDDLEHGSHV 28
 |||||
 DB 463 HGHEQOHGIGHGKFKRDDLEHGSHV 490

RESULT 2
 C25486
 K-Kininogen, HMW precursor - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 30-Sep-1993
 C:Accession: C25486
 R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 262, 2190-2198, 1987
 A:Title: Differing expression patterns and evolution of the rat kininogen gene family
 A:Reference number: A92625; MUID:87137443
 A:Accession: C25486
 A:Molecule type: DNA
 A:Residues: 1-264 <KIT>
 C:Comment: The nucleotide sequence was obtained from GenBank, release 55.0.
 C:Superfamily: kininogen; cystatin homology

Query Match 59.0%; Score 98.5; DB 2; Length 264;
 Best Local Similarity 50.0%; Pred. No. 4.4e-06;
 Matches 18; Conservative 3; Mismatches 6; Indels 9; Gaps 1;

QY 1 HGHEQOHGIGHGKFKRDDLEHGSHV 27
 |||||
 DB 75 HGHEQOHGIGHGKFKRDDLEHGSHV 110


```
RESULT 3
A25486
K:kininogen, HMW I precursor - rat
N:Contains: bradykinin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C:Accession: A25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2196, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family.
A:Reference number: A92625; MUID:87137443
A:Accession: A25486
A:Molecule type: mRNA
A:Residues: 1-639 <KIT>
A>Note: The authors translated the codon CAA for residue 347 as Asn
C:Superfamily: kininogen; cystatin homology
C:Keywords: alternative splicing
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-639/Product: kininogen, HMW I #status predicted <MAT>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>

Query Match 59.0%; Score 98.5; DB 2; Length 639;
Best Local Similarity 50.0%; Pred. No. 1.1e-05;
Matches 18; Conservative 3; Mismatches 6; Indels 9; Gaps 1;

QY 1 HGHEOHHGLGHGKFKRLD-----DLEHOGGH 27
      |||: ||||| ||| ||| |||
DB 450 HGOKPHGLGHGKFKRLDLDLKOORDDGYVYRHPMGH 485

RESULT 4
C27115
K:kininogen, LMW precursor - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
C:Accession: C27115; A25488
R:Fung, W.P.; Schreiber, G.
J. Biol. Chem. 262, 9298-9308, 1987
A:Title: Structure and expression of the genes for major acute phase alpha-1-protein (th
A:Reference number: A92653; MUID:87250580
A:Accession: C27115
A:Molecule type: DNA
A:Residues: 1-290 <FUN>
R:Kageyama, R.; Kitamura, N.; Okubo, H.; Nakanishi, S.
J. Biol. Chem. 262, 2345-2351, 1987
A:Title: Differing utilization of homologous transcription initiation sites of rat K and
A:Reference number: A25488; MUID:87137465
A:Accession: A25488
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-48 <RAG>
A:Cross-references: GB:J02662; NID:g205071; PIDN:AAA4183.1; PID:g205072
C:Superfamily: kininogen; cystatin homology
F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 54.8%; Score 91.5; DB 2; Length 290;
Best Local Similarity 47.2%; Pred. No. 4.3e-05;
Matches 17; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

QY 1 HGHEOHHGLGHGKFKRLD-----DLEHOGGH 27
      |||: ||||| ||| ||| |||
DB 124 NGOKPHGLGHGKFKRLDLDLKOORDDGYVYRHPMGH 159

RESULT 5
A27115
major acute phase alpha-1 protein 1 - rat (fragments)
```

```
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A27115
R:Fung, W.P.; Schreiber, G.
J. Biol. Chem. 262, 9298-9308, 1987
A:Title: Structure and expression of the genes for major acute phase alpha-1-protein
A:Reference number: A92653; MUID:87250580
A:Accession: A27115
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-315 <FUN>
C:Genetics:
A:Gene: MAP1
C:Superfamily: kininogen; cystatin homology
F:19-65/Domain: cystatin homology (fragment) <CYS>

Query Match 54.8%; Score 91.5; DB 2; Length 315;
Best Local Similarity 47.2%; Pred. No. 4.7e-05;
Matches 17; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

QY 1 HGHEOHHGLGHGKFKRLD-----DLEHOGGH 27
      |||: ||||| ||| ||| |||
DB 148 NGOKPHGLGHGKFKRLDLDLKOORDDGYVYRHPMGH 183

RESULT 6
KGB0H2
K:kininogen, HMW II precursor - bovine
N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C:Species: Bos primigenius taurus (cattle)
C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C:Accession: A01282; A91923; A91941; A91938; B29559
R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A:Title: A single gene for bovine high molecular weight and low molecular weight kini
A:Reference number: A93317; MUID:84014106
A:Accession: A01282
A:Molecule type: mRNA
A:Residues: 1-619 <KIT>
A:Cross-references: GB:V01492; GB:K01758; NID:g493; PIDN:CAA24736.1; PID:g494
R:Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
A:Reference number: A91923; MUID:70180420
A:Accession: A91923
A:Molecule type: protein
A:Residues: 376-391 <KAT>
R:Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.
J. Biochem. 79, 1201-1222, 1976
A:Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amin
A:Reference number: A91941; MUID:76260155
A:Accession: A91941
A:Molecule type: protein
A:Residues: 387-455 <HAN>
A>Note: 398-Pro, 401-Val, and 455-Iys were also found
R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen.
A:Reference number: A91938; MUID:75170265
A:Accession: A91938
A:Molecule type: protein
A:Residues: 456-496 <HA2>
R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwana
J. Biol. Chem. 262, 2768-2779, 1987
A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions o
A:Reference number: A92627; MUID:87137530
A:Accession: B29559
A:Molecule type: protein
A:Residues: 'Z', '20-104', 'E', '106-256', 'XX', '257-376 <SUB>
R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foerisch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
```

A>Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen.
 A:Reference number: A91153; MUID:86030270
 A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R:Seiyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 S:Seiyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 A>Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation; disulfide bonds
 A:Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is implicated in the release of bradykinin.
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-619/Product: HMW kininogen II #status predicted <MAT>
 F:19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
 F:19-130/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:261-372/Domain: cystatin homology <CY3>
 F:378-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:378-619/Product: HMW kininogen II light chain #status experimental <LCH>
 F:418-488/Region: glycine/histidine/lysine-rich
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
 F:27-589, 82-93, 106-125, 141-144, 205-217, 228-247, 261-264, 325-337, 348-361/Disulfide bonds:
 F:47/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:77, 168, 169, 204, 280/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
 F:197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:380/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:396, 400, 404, 510/Binding site: carbohydrate (Ser) (covalent) #status experimental
 F:397, 398, 518, 522, 534, 546, 551, 568/Binding site: carbohydrate (Thr) (covalent) #status ex
 F:496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 50.9%; Score 85; DB 1; Length 619;
 Best Local Similarity 55.6%; Pred. No. 0.00073;
 Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

OY 1 HGHEDQHGHCGRKFLDDLEHGSH 27
 |||:|||||||
 Db 461 HGHQKHGHCGRKFLDDLEHGSH 479

RESULT 7
 KGBOH1
 Kininogen, HMW I precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; prokininogen
 M:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A01281; A91923; A91938; A29559
 R:Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakashima, S.
 S:Nature 305, 545-549, 1983
 A>Title: A single gene for bovine high molecular weight and low molecular weight kininogen
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01281
 A:Molecule type: mRNA
 A:Residues: 1-621 <KIT>
 A:Cross-references: GB:V01491; GB:K01757; NID:9491; PIDN:CAA24735.1; PID:9492
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A>Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 378-393 <KAT>
 R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.

J. Biochem. 77, 55-68, 1975
 A>Title: Studies on the primary structure of bovine high-molecular-weight kininogen.
 A:Reference number: A91938; MUID:75170265
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 458-498 <HAN>
 R:Seiyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwana
 J. Biol. Chem. 262, 2768-2779, 1987
 A>Title: Bovine high molecular weight kininogen. The amino acid sequence, positions o
 A:Reference number: A92627; MUID:87137530
 A:Accession: A29559
 A:Molecule type: protein
 A:Residues: 1-20-123, 125-127, 129-378 <SUE>
 R:Kotisch, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A>Title: The amino acid sequence of the light chain of human high-molecular-mass kint
 A:Reference number: A91153; MUID:86030270
 A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R:Seiyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 S:Seiyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 A>Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation; disulfide bonds
 A:Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is implicated in the release of bradykinin.
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-621/Product: HMW prokininogen I #status predicted <MAT>
 F:19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
 F:19-130/Domain: cystatin homology <CY1>
 F:141-252/Domain: cystatin homology <CY2>
 F:263-374/Domain: cystatin homology <CY3>
 F:379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
 F:380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
 F:417-488/Region: glycine/histidine/lysine-rich
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:27-591, 82-93, 106-125, 141-144, 205-217, 228-247, 263-266, 327-333, 350-369/Disulfide bond
 F:87, 168, 169, 204/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
 F:197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:376-379/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:382/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:398, 406, 512/Binding site: carbohydrate (Ser) (covalent) #status experimental
 F:399, 400, 520, 524, 536, 548, 553, 570/Binding site: carbohydrate (Thr) (covalent) #status
 F:496-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 50.9%; Score 85; DB 1; Length 621;
 Best Local Similarity 55.6%; Pred. No. 0.00073;
 Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

OY 1 HGHEDQHGHCGRKFLDDLEHGSH 27
 |||:|||||||
 Db 463 HGHQKHGHCGRKFLDDLEHGSH 481

RESULT 8
 A33742
 female sterile homeotic protein, 205k - fruit fly (Drosophila melanogaster)
 N:Alternate names: membrane protein fsh, 205k
 M:Contains: female sterile homeotic protein, 110k
 C:Species: Drosophila melanogaster
 C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 20-Sep-1999
 C:Accession: A43742; B43742
 R:Haynes, S.R.; Mozer, B.A.; Bhalla-Dey, N.; Dawid, I.B.
 Dev. Biol. 134, 246-257, 1989

A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62474.1; PID:g1149500
A:Accession: S66150
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 'M',428-1085 <ME2>
A:Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62475.1; PID:g1149501
R:Horowitz, H.; Berg, C.A. Development 122, 1859-1871, 1996
A:Title: The Drosophila pipsqueak gene encodes a nuclear BTB-domain-containing protein
A:Reference number: Z22972; MUID:96232300
A:Accession: I45461
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355, 'E',357-1005, 'H',1007-1020, 'Q',1021-1061, 'ERS' <HOR>
A:Cross-references: EMBL:U48358; NID:g1203906; PIDN:AAC47153.1; PID:g1203907
A:Experimental source: tissue type ovarian
C:Genetics:
A:Gene: pipsqueak; psq
A:Map position: II
A:Introns: 427/3
C:Function:
A:Description: required for establishing polarity of the developing egg chamber
C:Superfamily: POZ domain homology
C:21-123/Domain: POZ domain homology <POZ>

Query Match	37.4%	Score 62.5;	DB 2;	Length 1085;
Best Local Similarity	44.4%;	Pred. No. 1.5;		
Matches 12; Conservative	0;	Mismatches 4;	Indels 11;	Gaps 1

QY	1	HGHEQDQGLGNGKFKLDDDLHQGGH	27
Db	332	HEHENHGCHGH-----GGH	347

RESULT 13
T35241
hypothetical protein SC5C7.34 SC5C7.34 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T35241
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21572
A:Accession: T35241
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-314 <SEE>
A:Cross-references: EMBL:AL031515; PDB:1CA20646.1; GSPDB:GN00070; SC5C7.34
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SC5C7.34

Query Match	37.1%;	Score 62;	DB 2;	Length 314;
Best Local Similarity	76.9%;	Pred. No. 0.47;		
Matches 10;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

OY	1	HGHEQHGHLGH	13
	1111	11111111	
Db	18	HGHGQGHGHGH	30

RESULT 14
T48099
hypothetical protein T20010.200 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cross)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000
C:Accession: T48099
R:Oermann, B.; Ottenwelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Lem
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24484
A:Accession: T48099

A:Status:preliminary
A:Molecule type: DNA
A:Residues: 1-199 <ORF>
A:Cross-references: EMBL:AL163816
A:Experimental source: cultivar Columba; BAC clone T20010
A:Genetics:
A:Map position: 3
A:Introns: 163/2
A:Note: T20010..200

Query Match	36.8%;	Score 61.5;	DB 2;	Length 199;
Best Local Similarity	38.2%;	Pred. No. 0.34;		
Matches 13;	Conservative 2;	Mismatches 12;	Indels 7;	Gaps 1

QY 1 HGEHQHGLGHGKFLDD-----DLEHOGG 27
||| : || ||| : || |||
Db 92 HGHGRGHGHGHGHRHGRDRHGRDRGHRHGRH 125

RESULT	15
T45059	

C:\Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:\Accession: T45059

B., O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, Nature 368, 32-38, 1994

A: Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*
A: Reference number: S43531; MUID: 94150718

A:Residues: 1-735 <Wtl>
A: Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DBJ

A: Experimental source: clone Y39B6
C: Genetics:
a. Man position: 3

A;Note: Y39B6B.gg

Best Local Similarity 40.7%; Pred. No. 1.6;
Matches 11; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

Oy	1	HGNEQDNGLGNHGKFKLDLDDLENGGN	27
		: :	
Db	516	HGNGENGSTNNGINGENNNARANNGN	542

Search completed: July 6, 2001, 09:18:01
Job time: 647 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:26:38 ; Search time 37.59 Seconds

(without alignments)
25.516 Million cell updates/sec

Title: US-09-437-912-5

Perfect score: 167
Sequence: 1 HGHEQHGHLGHGKFKLDLLEHGCHV 28

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	167	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	98.5	59.0	639	1 KNG_RAT	P08934 rattus norv
3	97.5	58.4	661	1 KNG_MOUSE	O08677 mus musculu
4	85	50.9	619	1 KNH2_BOVIN	P01045 bos taurus
5	85	50.9	621	1 KNH1_BOVIN	P01044 bos taurus
6	71	42.5	693	1 CAUP_DROME	P34269 drosophila
7	65	38.9	2038	1 FSH_DROME	P13709 drosophila
8	63.5	38.0	476	1 KE4_MOUSE	O31125 mus musculu
9	62.5	37.4	449	1 CSUP_DROME	O93344 drosophila
10	60.5	36.2	515	1 KEAL_CAEEL	O9xtq7 caenorhabdi
11	59	35.3	118	1 S109_RABIT	P50117 oryctolagus
12	58.5	35.0	469	1 KE4_HUMAN	P02504 homo sapien
13	57	34.1	305	1 HYPB_BRAJA	O45257 bradyrhizob
14	57	34.1	332	1 HYPB_RHCCA	P26410 rhodopacter
15	57	34.1	337	1 ZNPA_HAEIN	P44526 haemophilus
16	57	34.1	670	1 VGS0_HSV1	O00130 ictaluriid h
17	56	33.5	439	1 COR1_YEAST	P32798 saccharomyc
18	55.5	33.2	336	1 FI1A_MOUSE	P11088 mus musculu
19	55	32.9	352	1 KE4_BRARE	O9pub8 brachydanio
20	53.5	32.0	258	1 BOX5_NOTVI	P53771 notophtalm
21	53	31.7	85	1 ANFE_SARPE	O08617 sarcophaga
22	53	31.7	1061	1 TRC4_ECOLI	P27189 escherichia
23	53	31.7	1448	1 TRC5_ECOLI	P27190 escherichia
24	52.5	31.4	584	1 D39B_ARATH	O06737 arabidopsis
25	52	31.1	503	1 ZNT1_MOUSE	O60738 mus musculu
26	52	31.1	507	1 ZNT1_RAT	O62720 rattus norv
27	51.5	30.8	212	1 STVD_AERHY	O07046 aeromonas h
28	51.5	30.8	527	1 HSR8_LYPE	P41153 lycopersico
29	51	30.5	306	1 CH38_DROME	P07183 drosophila
30	51	30.5	496	1 BAP1_KLUMA	P33293 kluyveromyc
31	50	29.9	382	1 PO3A_XENLA	P31365 xenopus lae
32	50	29.9	623	1 Z255_HUMAN	O9uid9 homo sapien
33	50	29.9	1174	1 PTRNL_HUMAN	O16825 homo sapien

34	49.5	29.6	414	1 Y288_MYCGE	P47530 mycoplasma
35	49	29.3	108	1 AGN3_APICA	P01364 aplysia cal
36	49	29.3	419	1 GSC_DROME	P54366 drosophila
37	48	28.7	59	1 HPM_HELPY	O48251 helicobacte
38	48	28.7	139	1 SALA_DROST	P21749 drosophila
39	48	28.7	142	1 SALA_DROME	P21750 drosophila
40	48	28.7	142	1 SALA_DROOR	P21748 drosophila
41	48	28.7	382	1 YHH6_YEAST	P32793 saccharomyc
42	48	28.7	510	1 FRK_DROME	P14734 drosophila
43	48	28.7	562	1 TPPI_MOUSE	O89023 mus musculu
44	48	28.7	601	1 PDML_DROME	P31368 drosophila
45	48	28.7	2870	1 NDVB_RHIME	P20471 rhizobium m

ALIGNMENTS

RESULT 1
KNG_HUMAN STANDARD: PRT: 644 AA.
ID P01042: P01043: 21-JUL-1986 (Rel. 01, Created)
AC 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KININOGEN PRECURSOR (ALPHA-2-THIOL PROTEINASE INHIBITOR) [CONTAINS: DE BRADYKININ].
CN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular weight and low molecular weight prekininogens. Primary structures of two human prekininogens.";
RT J. Biol. Chem. 260:8601-8609(1985).
RL [2]
RN GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T., Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for its evolution.";
RT J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiohara H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and its identity with low molecular weight kininogen.";
RT Biochemistry 23:5691-5697(1984).
RN [4]
RP SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foerisch B., Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-mass kininogen.";
RT Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RT Fed. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

RL Selkagaku 56:808-808(1984).

CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES: (2)

CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY

CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO

CC FACTOR XIII: (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-

CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE

CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF

CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE

CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND

CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A

CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR

CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF

CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS

CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,

CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)

CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-

CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD

CC CLOTTING.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>

CC or send an email to license@slb-slb.ch).

CC -----

DR EMBL: K02566: AAA35497.1: -

DR EMBL: M11437: AAB59550.1: -

DR EMBL: M11438: AAB59550.1: JOINED.

DR EMBL: M11521: AAB59550.1: JOINED.

DR EMBL: M11522: AAB59550.1: JOINED.

DR EMBL: M11523: AAB59550.1: JOINED.

DR EMBL: M11524: AAB59550.1: JOINED.

DR EMBL: M11525: AAB59550.1: JOINED.

DR EMBL: M11526: AAB59550.1: JOINED.

DR EMBL: M11527: AAB59550.1: JOINED.

DR EMBL: M11528: AAB59550.1: JOINED.

DR EMBL: M11437: AAB59551.1: -

DR EMBL: M11438: AAB59551.1: JOINED.

DR EMBL: M11521: AAB59551.1: JOINED.

DR EMBL: M11522: AAB59551.1: JOINED.

DR EMBL: M11523: AAB59551.1: JOINED.

DR EMBL: M11524: AAB59551.1: JOINED.

DR EMBL: M11525: AAB59551.1: JOINED.

DR EMBL: M11526: AAB59551.1: JOINED.

DR EMBL: M11527: AAB59551.1: JOINED.

DR EMBL: M11528: AAB59551.1: JOINED.

DR PIR: A01279: KGHUHL.

DR PIR: A25276: A25276.

DR PIR: A01280: KGHUHL.

DR PIR: B25276: B25276.

DR PIR: S02482: S02482.

DR SWISS-2DPAGE: P01043: HUMAN.

DR MIM: 228960: -

DR InterPro: IPR000010: -

DR InterPro: IPR002395: -

DR Pfam: PF00031: cystatin.3.

DR PRINTS: PR00334: KININOGEN.

DR PROSITE: PS00287: CYSTATIN.2.

DR Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;

DR Bradykinin; Blood coagulation; Inflammatory response; Signal;

KW Alternative splicing.

FT SIGNAL 1 18

FT CHAIN 19 644 KININOGEN,

FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 389 BRADYKININ.

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD_RES	19	19	PYRROLIDONE CARBOXYLIC ACID.
FT	DISULFID	28	614	INTERCHAIN.
FT	DISULFID	83	94	
FT	DISULFID	107	126	
FT	DISULFID	142	145	
FT	DISULFID	206	218	
FT	DISULFID	229	248	
FT	DISULFID	264	267	
FT	DISULFID	328	340	
FT	DISULFID	351	370	
FT	CARBOHYD	48	48	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	205	205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	294	294	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	401	401	
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPPLIC	402	427	VSPPTSMAPADDERDSKDEGHR -> SHLRSCFEKGR
FT	VARSPPLIC	428	644	PKKABRPASREREVS (IN ISOFORM LMW).
FT	CONFLICT	593	593	MISSING (IN ISOFORM LMW).
FT	SEQUENCE	644 AA: 71945 MW: 313284	644 AA: 71945 MW: 313284	CBABF8FB7E CRC64;

Query Match 100.0%; Score 167; DR 1; Length 644;

Best Local Similarity 100.0%; Pred. No. 2.9e-15;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HGHEQHGGLGHGKFKRLDDLEHOGSHV 28

Db 463 HGHEQHGGLGHGKFKRLDDLEHOGSHV 490

RESULT 2

ID	KNG_RAT	STANDARD:	PRT:	639 AA.
AC	P08934: P08933:			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-NOV-1988 (Rel. 09, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	KININOGEN PRECURSOR [CONTAINS: BRADYKININ].			
GN	KNG.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RP	SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).			
RP	MEDLINE=87137443; PubMed=3029068;			
RA	Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;			
RT	"Differing expression patterns and evolution of the rat kininogen			
RL	gene family.";			
RN	J. Biol. Chem. 262:2190-2198(1987).			
RP	SEQUENCE FROM N.A. (LMW ISOFORM).			
RX	MEDLINE=86008264; PubMed=2413018;			
RA	Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;			
RT	"Primary structures of the mRNAs encoding the rat precursors for			

RT bradykinin and T-kinin. Structural relationship of kininogens with
 RT major acute phase protein and alpha 1-cysteine proteinase
 RL inhibitor. J. Biol. Chem. 260:12054-12059(1985).
 RN [3]
 RC SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN-BUFFALO;
 RX MEDLINE=87250580; PubMed=2439509;
 RA Fung W.-P., Schreiber G.;
 RT "Structure and expression of the genes for major acute phase alpha 1-
 RL protein (thioesterin) and kininogen in the rat";
 RL J. Biol. Chem. 262:9298-9308(1987).
 RN [4]
 RP SEQUENCE OF 1-41 FROM N.A.
 RP STRAIN-WISTAR; TISSUE=Liver;
 RX MEDLINE=87137465; PubMed=3818598;
 RA Kagayama R., Kitamura N., Ohkubo H., Nakanishi S.;
 RT "Differing utilization of homologous transcription initiation sites
 of rat K and T kininogen genes under inflammation condition.";
 RL J. Biol. Chem. 262:2345-2351(1987).
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
 CC HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
 CC T-II.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L29428; AAA41486.1; -
 CC EMBL: M11884; AAA41487.1; -
 CC EMBL: M14369; AAA41484.1; -
 CC EMBL: M14369; AAA41485.1; ALT-SEQ.
 CC EMBL: M16455; AAA41482.1; -
 CC PIR: A25486; A25486.
 CC PIR: A28055; A28055.
 CC InterPro: IPR000010; -
 CC InterPro: IPR002395; -
 CC Pfam: PF00031; cystatin.3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSTATIN.2.
 CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 CC Bradykinin; Blood coagulation; Inflammatory response; Signal;
 CC Alternative splicing; Multigene family.
 FT CHAIN 1 18 KININOGEN.
 FT SIGNAL 19 639 KININOGEN.
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.

FT PEPTIDE 381 389 BRADYKININ.
 FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 FT DOMAIN 439 514 HIS-RICH.
 FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 107 126 BY SIMILARITY.
 FT DISULFID 142 145 BY SIMILARITY.
 FT DISULFID 206 218 BY SIMILARITY.
 FT DISULFID 229 248 BY SIMILARITY.
 FT DISULFID 264 267 BY SIMILARITY.
 FT DISULFID 328 340 BY SIMILARITY.
 FT DISULFID 351 370 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 284 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 402 433 VSPSYIAVQERDGNQGFHGWLAHAR -> RLINS
 CEYKRLKAGGAPAPERQAASTVTP (IN ISOFORM
 LMW).
 FT VARSPLIC 434 639 MISSING (IN ISOFORM LMW).
 FT CONFLICT 61 61 E -> K (IN REF. 2).
 FT SEQUENCE 639 AA: 70933 MM: D3172DF94FF56AF5 CRC64;
 SQ
 Query Match 59.0%; Score 98.5; DB 1; Length 639;
 Best Local Similarity 50.0%; Pred. No. 5, 1e-06;
 Matches 18; Conservative 3; Mismatches 6; Indels 9; Gaps 1;
 QY 1 HGHEDQHGSGHGRKRLD-----DLEHQGGH 27
 DB 450 HGHQKPHGSGHGHQKRLDLDKQREDGYDHRPVGH 485
 RESULT 3
 KNG_MOUSE STANDARD: PRT; 661 AA.
 ID KNG_MOUSE 008677; 008676;
 AC 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
 GN KNG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxId=10090;
 RX SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RP STRAIN=C57BL/6 x CBA; TISSUE=Liver;
 RA Takano M., Kondoh J., Yamaoka K., Okamoto H.;
 RT "Molecular cloning of cDNAs for mouse low- and high- molecular
 kininogen.";
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-

CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC CLOTTING (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D84435; BAA19743.1; -
CC EMBL: D84415; BAA19742.1; -
CC MGI: 1097705; Kng.
CC InterPro: IPR000010; -
CC InterPro: IPR002395; -
CC InterPro: IPR003243; -
CC Pfam: PF00031; Cystatin: 3.
CC PRINTS: PRO0334; KININOGEN.
CC PROSITE: PS00287; CYSTATIN; 1.
CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
CC Bradykinin; Blood coagulation; Inflammatory response; Signal;
CC Alternative splicing.
CC FT SIGNAL 1 18
CC FT CHAIN 19 661
CC FT CHAIN 19 379
CC FT CHAIN 380 388
CC FT CHAIN 389 661
CC FT DOMAIN 19 135
CC FT DOMAIN 136 257
CC FT DOMAIN 238 379
CC FT DOMAIN 439 524
CC FT DISULFID 28 631
CC FT DISULFID 83 94
CC FT DISULFID 107 125
CC FT DISULFID 141 144
CC FT DISULFID 205 217
CC FT DISULFID 228 247
CC FT DISULFID 263 266
CC FT DISULFID 327 339
CC FT CARBOHYD 350 369
CC FT CARBOHYD 82 82
CC FT CARBOHYD 168 168
CC FT CARBOHYD 204 204
CC FT CARBOHYD 242 242
CC FT VARSPLIC 401 432
CC FT VARSPLIC 433 661
CC FT SEQUENCE 661 AA; 73102 MW; 774460258D587966 CRC64;
CC
CC Query Match 58.4%; Score 97.5; DB 1; Length 661;
CC Best Local Similarity 50.0%; Pred. No. 7.3e-06;
CC Matches 18; Conservative 3; Mismatches 6; Indels 9; Gaps 1;
CC
CC OY 1 HGHEQDQHGHCHEKFLD-----DDLEHGGH 27
CC |||: |||||: |||
CC Db 460 HGHEQDQHGHCHEKFLDYLRLHQRGEGDDHTHYGH 495
CC
CC RESULT 4
CC KHM2_BOVIN KHM2_BOVIN STANDARD: PRT; 619 AA.
CC AC P01045:
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KININOGEN, HMW II PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS:
DE BRADYKININ].
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84014106; PubMed=6571699;
RA Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT "A single gene for bovine high molecular weight and low molecular
RT weight kininogens.";
RL Nature 305:545-549 (1983).
RN [2]
RP SEQUENCE OF 19-376.
RX MEDLINE=87137530; PubMed=3546295;
RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA Miyata T., Iwanaga S.;
RT "Bovine high molecular weight kininogen. The amino acid sequence,
RT positions of carbohydrate chains and disulfide bridges in the heavy
RT chain portion.";
RL J. Biol. Chem. 262:2768-2779 (1987).
RN [3]
RP SEQUENCE OF 376-391.
RX MEDLINE=70180420; PubMed=4986212;
RA Kato H., Nagasawa S., Suzuki T.;
RT "Studies on the structure of bovine kininogen: cleavages of disulfide
RT bonds and of methionyl bonds in kininogen-II.";
RL J. Biochem. 67:313-323 (1970).
RN [4]
RP SEQUENCE OF 387-455.
RX MEDLINE=76260155; PubMed=956151;
RA Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT "Primary structure of bovine plasma high-molecular-weight kininogen.
RT The amino acid sequence of a glycopeptide portion (fragment 1)
RT following the C-terminus of the bradykinin moiety.";
RL J. Biochem. 79:1201-1222 (1976).
RN [5]
RP SEQUENCE OF 456-496.
RX MEDLINE=75170265; PubMed=1169237;
RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT "Studies on the primary structure of bovine high-molecular-weight
RT kininogen. Amino acid sequence of a fragment ('histidine-rich
RT peptide') released by plasma kallikrein.";
RL J. Biochem. 77:55-68 (1975).
CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES: (2)
CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC TO FACTOR XII: (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES: (4) THE ACTIVE
CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC NARITRESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC ACTION).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC TO RESIDUE 398.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: V01492; CAA24736.1; ALT_SEQ.

DR PIR: A01282; KGBOH2.

DR PIR: B28559; B29559.

DR HSSP: P04129; IAFI.

DR InterPro: IPR000010; .

DR InterPro: IPR002395; .

DR Pfam: PF00031; cystatin. 3.

DR PRINTS: PR00334; KININOGEN.

DR PROSITE: PS00287; CYSTATIN; 2.

KW Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;

KW Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;

KW Inflammatory response.

FT SIGNAL 1 18

FT CHAIN 19 619 KININOGEN, HMW II.

FT CHAIN 19 376 HEAVY CHAIN.

FT PEPTIDE 378 386 BRADYKININ.

FT CHAIN 387 619 LIGHT CHAIN.

FT DOMAIN 19 135 CYPSTATIN-LIKE 1.

FT DOMAIN 136 256 CYPSTATIN-LIKE 3.

FT DOMAIN 257 376 CYPSTATIN-LIKE 2.

FT MOD_RES 19 19 PYRROLIDONE CARBOXYLIC ACID.

FT CARBOHYD 136 87 N-LINKED (GLCNAC. . .).

FT CARBOHYD 168 136 N-LINKED (GLCNAC. . .).

FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).

FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).

FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).

FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .).

FT DISULFID 27 589 INTERCHAIN.

FT DISULFID 82 93

FT DISULFID 106 125

FT DISULFID 141 144

FT DISULFID 205 217

FT DISULFID 228 247

FT DISULFID 261 264

FT DISULFID 325 337

FT DISULFID 348 367

FT VARIANT 398 398 T -> P.

FT VARIANT 401 401 L -> V.

FT VARIANT 454 454 H -> K.

SEQUENCE 619 AA: 68710 MW: F04320A8EB0EB0DA CRC64:

Query Match Score 85; DB 1; Length 619;
Best Local Similarity 55.6%; Pred. No. 0.00033;
Matches 15; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

OY 1 HGHEOQHGCHGKFLDDLEHGGH 27
DB 461 HGKQKHGHLGHK-----HGCH 479

RESULT 5
ID KNL_BOVIN STANDARD: PRT: 621 AA.

AC P01044;

DT 21-JUL-1986 (rel. 01, Created)

DT 21-JUL-1986 (rel. 01, Last sequence update)

DT 01-JUN-1994 (rel. 29, Last annotation update)

DE KININOGEN, HMW I PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS:

DE BRADYKININ].

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE-84014106; PubMed-6571699;
RT Kitamura N., Takagaki Y., Furuta S., Tanaka T., Nawa H., Nakanishi S.,
RT "A single gene for bovine high molecular weight and low molecular
RT weight kininogens.";
RL Nature 305:545-549(1983).

RN [2]
RX SEQUENCE OF 19-378.
RA MEDLINE-87137530; PubMed-3546295;
RT Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA Miyata T., Iwanaga S.;
RT "Bovine high molecular weight kininogen. The amino acid sequence,
RT positions of carbohydrate chains and disulfide bridges in the heavy
RT chain portion.";
RL J. Biol. Chem. 262:2768-2779(1987).

RN [3]
RX SEQUENCE OF 378-393.
RA MEDLINE-70180420; PubMed-4986212;
RT Kato H., Nagasawa S., Suzuki T.;
RT "Studies on the structure of bovine kininogen: cleavages of disulfide
RT bonds and of methionyl bonds in kininogen-II.";
RL J. Biochem. 67:313-323(1970).

RN [4]
RX SEQUENCE OF 458-498.
RA MEDLINE-75170265; PubMed-1169237;
RT Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT "Studies on the primary structure of bovine high-molecular-weight
RT kininogen. Amino acid sequence of a fragment ('histidine-rich
RT peptide') released by plasma kallikrein.";
RL J. Biochem. 77:55-68(1975).

CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE
CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC MUSCLE CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C)
CC NATURESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL,
CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC ACTION).

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- ALTERNATIVE PRODUCTS: HMW I AND LMW I KININOGEN PRECURSORS ARE
CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC TO RESIDUE 400.

CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC -1- SIMILARITY: KININOGEN IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: V01491; CAA24735.1; .

CC PIR: A01281; KGBOH1.

CC PIR: A29559; A29559.

CC InterPro: IPR000010; .

CC InterPro: IPR002395; .

CC Pfam: PF00031; cystatin. 3.

CC PRINTS: PR00334; KININOGEN.

CC PROSITE: PS00287; CYPSTATIN; 2.

CC Glycoprotein; Plasma; Duplication; Vasodilator; Alternative splicing;

CC Thiol protease inhibitor; Bradykinin; Blood coagulation;

KW Inflammatory response; Signal.
 FT SIGNAL 1 18 PROBABLE.
 FT CHAIN 19 621 KININOGEN, HMW I.
 FT CHAIN 19 378 BEADY CHAIN.
 FT PEPTIDE 380 388 BRADYKININ.
 FT CHAIN 389 621 LIGHT CHAIN.
 FT DOMAIN 19 135 CYSTATIN-LIKE 1.
 FT DOMAIN 136 257 CYSTATIN-LIKE 2.
 FT DOMAIN 258 378 CYSTATIN-LIKE 3.
 FT MOD_RES 19 19 PYROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 136 87 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 136 136 PARTIAL.
 FT CARBOHYD 168 168 OR 169.
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
 FT DISULFID 27 591 INTERCHAIN.
 FT DISULFID 82 93
 FT DISULFID 106 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 SQ SEQUENCE 621 AA; 68890 MW; D16850BEFE3C55CD CRC64;

Query Match 50.9%; Score 85; DB 1; Length 621;
 Best Local Similarity 55.6%; Pred. No. 0.00033; Mismatches 2; Indels 8; Gaps 1;
 Matches 15; Conservative 2;

QY 1 HGHEQOHGLGHGKFKLDDLEHOGCH 27
 DB 463 HGHQOHGLGHGK-----HGHH 481

RESULT 6
 CAUP_DROME STANDARD; PRT; 693 AA.
 AC P54269;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HOMEOBOX PROTEIN CAUPOLICAN.
 GN CAUP.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Ephydroidea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96180722; PubMed=8620542;
 RA Gomez-Skarneta J.-L., del Corral R.D., de la Calle-Mustienes E.,
 RA Ferrer-Marco D., Modolell J.,
 RT "Araucan and caupolican, two members of the novel iroquois complex,
 RT encode homeoproteins that control proneural and vein-forming genes.";
 RT Cell 85:95-110(1996).
 RL -1- FUNCTION: CONTROLS PRONEURAL AND VEIN FORMING GENES. POSITIVE
 CC TRANSCRIPTIONAL CONTROLLER OF AC-SC (ACHAETE-SCUTE). MAY ACT AS AN
 CC ACTIVATOR THAT INTERACTS WITH THE TRANSCRIPTIONAL COMPLEX
 CC ASSEMBLED ON THE AC AND SC PROMOTERS AND PARTICIPATES IN
 CC TRANSCRIPTION INITIATION.
 CC -1- SIMILARITY: BELONGS TO THE TALE/TRO FAMILY OF HOMEOBOX PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X95178; CAA64485.1; -
 DR HSSP: P02833; ISAN.
 DR FlyBase: FBgn0015919; caup.
 DR InterPro: IPR001356; -.
 DR Pfam: PF00046; homeobox; 1.
 DR PROSITE: PS00027; homeobox; 1.
 DR PROSITE: PS00071; HOMEOBOX_2; 1.
 KW Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
 KW Developmental protein.
 FT DNA_BIND 226 288 HOMEOBOX (TALE-TYPE).
 FT DOMAIN 300 303 POLY-ASP.
 FT DOMAIN 405 418 POLY-GLN.
 FT DOMAIN 501 516 POLY-GLN.
 FT DOMAIN 517 528 POLY-HIS.
 FT DOMAIN 565 572 POLY-SER.
 FT DOMAIN 613 624 POLY-SER.
 SQ SEQUENCE 693 AA; 73749 MW; 8E0D6D43C9C6D619 CRC64;

Query Match 42.5%; Score 71; DB 1; Length 693;
 Best Local Similarity 50.0%; Pred. No. 0.029; Mismatches 7; Indels 6; Gaps 1;
 Matches 14; Conservative 1;

QY 1 HGHEQOHGLGHGKFKLDDLEHOGCHV 28
 DB 656 HGHHGHGHLGHG-----GLGHGHGM 677

RESULT 7
 FSH_DROME STANDARD; PRT; 2038 AA.
 AC P13709; P13710;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE FEMALE STERILE HOMEOBOX PROTEIN (FRAGILE-CHORION MEMBRANE PROTEIN).
 GN FS(1)H OR FSH.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Ephydroidea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89276730; PubMed=2567251;
 RA Haynes S.R., Mozer B.A., Bhatia-Dey N., David I.B.,
 RT "The Drosophila fsh locus, a maternal effect homeotic gene, encodes
 RT apparent membrane proteins.";
 RT Dev. Biol. 134:246-257(1989).
 CC -1- FUNCTION: REQUIRED MATERIALLY FOR PROPER EXPRESSION OF OTHER
 CC HOMEOBOX GENES INVOLVED IN PATTERN FORMATION, SUCH AS UBX.
 CC
 CC -1- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.
 CC
 CC -1- SIMILARITY: CONTAINS 2 BROWDOMAINS.
 CC
 CC -1- SIMILARITY: CONTAINS 1 ET DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklow G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.).
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: NEGATIVELY REGULATES TYROSINE HYDROXYLASE ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF216584; AAF37226.1; -
 CC EMBL: AE003661; AAF53744.1; -
 CC EMBL: AF000002; Catsup.
 CC Transmembrane.
 CC KW TRANSMEM 19 39 POTENTIAL.
 CC FT TRANSMEM 135 155 POTENTIAL.
 CC FT TRANSMEM 167 187 POTENTIAL.
 CC FT TRANSMEM 222 242 POTENTIAL.
 CC FT TRANSMEM 371 391 POTENTIAL.
 CC FT TRANSMEM 395 415 POTENTIAL.
 CC FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SO SEQUENCE 449 AA; 48658 MW; F711A254C07AB4C CRC64;
 CC -----
 CC Query Match 37.4%; Score 62.5; DB 1; Length 449;
 CC Best Local Similarity 42.9%; Pred. No. 0.26; Mismatches 13; Indels 1; Gaps 1;
 CC Matches 12; Conservative 2; Mismatches 13; Indels 1; Gaps 1;
 CC -----
 CC QY 1 HGHEDQHGICGHRK-LDDLEHOGGH 27
 CC DB 83 HDHHDHGHGHHGHGHHGHGHHGHGHHGH 110
 CC -----
 CC RESULT 10
 CC KE4L.CAEEL STANDARD; PRT; 515 AA.
 CC AC 09X07:
 CC DT 01-OCT-2000 (Rel. 40, Created)
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
 CC DE HYPOTHETICAL KE4-LIKE PROTEIN H13N06.5 IN CHROMOSOME X.
 CC GN H13N06.5.
 CC OS Caenorhabditis elegans.
 CC OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC OC Rhabditidae; Peloderinae; Caenorhabditis.
 CC OX NCBI_TaxID=6239;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-BRISTOL N2;
 CC RA Leonard N.;
 CC RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -----

CC -1- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z99942; CAB17070.1; -
 CC DR Wormpep; H13N06.5; CE18815.
 CC KW Hypothetical protein; Transmembrane; Glycoprotein.
 CC FT TRANSMEM 27 47 POTENTIAL.
 CC FT TRANSMEM 49 69 POTENTIAL.
 CC FT TRANSMEM 214 234 POTENTIAL.
 CC FT TRANSMEM 247 267 POTENTIAL.
 CC FT TRANSMEM 297 317 POTENTIAL.
 CC FT TRANSMEM 386 406 POTENTIAL.
 CC FT TRANSMEM 429 449 POTENTIAL.
 CC FT TRANSMEM 463 483 POTENTIAL.
 CC FT TRANSMEM 92 182 HIS-RICH.
 CC FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SO SEQUENCE 515 AA; 55500 MW; 17D7E854F4E1DAAF CRC64;
 CC -----
 CC Query Match 36.2%; Score 60.5; DB 1; Length 515;
 CC Best Local Similarity 39.3%; Pred. No. 0.55;
 CC Matches 11; Conservative 3; Mismatches 13; Indels 1; Gaps 1;
 CC -----
 CC QY 1 HGHEDQHGICGHRK-FKLDLEHOGGH 27
 CC DB 139 HGHHDHGHGHHGHGHHGHGHHGHGHHGH 166
 CC -----
 CC RESULT 11
 CC S109-RABIT STANDARD; PRT; 118 AA.
 CC AC P50117;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
 CC DE CALGRANTLIN B (MIGRATION INHIBITORY FACTOR-RELATED PROTEIN 14)
 CC DE (MRP-14) (FRAGMENT).
 CC GN S100A9 OR MRP-14.
 CC OS Oryctolagus cuniculus (Rabbit).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC OX NCBI_TaxID=9986;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-NEW ZEALAND WHITE; TISSUE-Neutrophils;
 CC RX MEDLINE=96355278; PubMed=8702688;
 CC RA Yang Z., de Veer M.J., Gardner E.E., Devanish R.J., Handley C.J.,
 CC RA Underwood J.R., Robinson H.C.;
 CC RT "Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-
 CC calgranulin C when incubated with inorganic [35S]sulfate.";
 CC RL J. Biol. Chem. 271:19802-19809(1996).
 CC RN [2]
 CC RP SEQUENCE OF 45-82 FROM N.A.
 CC RC STRAIN-NEW ZEALAND WHITE;
 CC RX MEDLINE=94198229; PubMed=8148323;
 CC RA Mori S., Goto K., Goto F., Mutakami K., Ohkawara S., Yoshinaga M.;
 CC RT "Dynamic changes in mRNA expression of neutrophils during the course
 CC of acute inflammation in rabbits.";
 CC RL Int. Immunol. 6:149-156(1994).
 CC CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF091849; AAC61771.1; -
 DR EMBL: D17404; BAA04227.1; -
 DR InterPro: IPR001751; -
 DR InterPro: IPR002048; -
 DR InterPro: IPR002395; -
 DR Pfam: PF01023; S_100; 1.
 DR Pfam: PF00036; ehand; 1.
 DR PRINTS: PR00334; KINNOGEN.
 DR PROSITE: PS00018; EF_HAND; 1.
 DR PROSITE: PS00303; S100_CABP; 1.
 KW Calcium-binding; Repeat.
 FT NON_TER 1
 FT CA_BIND 9 22 SITE I (LOW AFFINITY) (POTENTIAL).
 FT CA_BIND 53 64 SITE II (HIGH AFFINITY) (POTENTIAL).
 FT DOMAIN 103 118 2 X 8 AA TANDEM REPEATS OF G-H-G-H-
 G-H-S-H.
 FT REPEAT 103 110 1.
 FT REPEAT 111 118 2.
 SQ SEQUENCE 118 AA; 13292 MW; 7496118E21AD5C41 CRC64;

Query Match 35.3%; Score 59; DB 1; Length 118;
 Best Local Similarity 69.2%; Pred. No. 0.18;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HGHGQHGIGHGH 13
 111 1111111

DB 104 HGHGSHGHGH 116

RESULT 12
 KE4_HUMAN
 ID KE4_HUMAN STANDARD; PRT; 469 AA.
 AC Q92504; Q9Q1Q0;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HISTIDINE-RICH MEMBRANE PROTEIN KE4.
 GN HKE4 OR KINGS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=97001166; PubMed=8812499;
 RA Ando A., Kikuchi Y., Shigenari A., Kawata H., Okamoto N., Shina T.,
 RA Chen L., Ikemura T., Abe K., Kimura M., Inoko H.;
 RT "cDNA cloning of the human homologues of the mouse Ke4 and Ke6 genes
 RT at the centromeric end of the human MHC region.";
 RT Genomics 35:600-602(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Verara A., Lana I., Corella A., de Miguel C., Migliaccio M.,
 RA Encio I.;
 RT "Molecular cloning and characterization of the human KE4 gene and 5'
 RT flanking region.";
 RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tubby B.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC -1- TISSUE SPECIFICITY: MAJOR EXPRESSION IN PLACENTA, LUNG, KIDNEY
 CC AND PANCREAS.

CC -1- SIMILARITY: BELONGS TO THE KE4/CATSPF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: D82060; BAA11528.1; -
 DR EMBL: AF117221; AAD12305.1; -
 DR EMBL: AL031228; CAA20238.1; -
 DR MIM: 601416; -
 KW Transmembrane; Glycoprotein.
 FT TRANSMEM 10 30
 FT TRANSMEM 138 158 POTENTIAL.
 FT TRANSMEM 169 189 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT TRANSMEM 381 401 POTENTIAL.
 FT TRANSMEM 417 436 POTENTIAL.
 FT DOMAIN 30 114 HIS-RICH.
 FT DOMAIN 238 263 HIS-RICH.
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 7 7 A -> G (IN REF. 1 AND 2).
 FT CONFLICT 280 280 E -> G (IN REF. 1 AND 2).
 FT CONFLICT 376 376 S -> T (IN REF. 1 AND 2).
 FT CONFLICT 397 469 CALLTEGAVSGEISLAGAPGAVLPTAGGFIYATVAVLP
 ELLREASPLQSLGVILGALGVIMVLAHLE -> VPFSL
 KEQWTVKIQVYVYLAGSCHLLQVALST (IN REF. 1
 AND 2).
 SQ SEQUENCE 469 AA; 50118 MW; 6504A1EF5AA65B9 CRC64;

Query Match 35.0%; Score 58.5; DB 1; Length 469;
 Best Local Similarity 40.7%; Pred. No. 0.93;
 Matches 11; Conservative 1; Mismatches 10; Indels 5; Gaps 1;

OY 1 HGHGQHGIGHGKRLDDLEHGGH 27
 111 111 111 111

DB 73 HGHTHDHDHGHSH-----EDLHHGSH 94

RESULT 13
 HTPB_BRAJA
 ID HTPB_BRAJA STANDARD; PRT; 302 AA.
 AC Q45257;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HYDROGENASE EXPRESSION/FORMATION PROTEIN HTPB.
 GN HTPB.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Bradyrhizobium group; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=94137733; PubMed=8305450;
 RA Fu C., Maier R.J.;
 RT "Nucleotide sequences of two hydrogenase-related genes (hnpA and
 RT hnpB) from Bradyrhizobium japonicum, one of which (hnpB) encodes an
 RT extremely histidine-rich region and guanine nucleotide-binding
 RT domains.";
 RT Biochim. Biophys. Acta 1184:135-138(1994).
 CC -1- FUNCTION: MAY WORK IN THE MOBILIZATION OF NICKEL INTO HYDROGENASE
 CC -1- ENZYME: BELONGS 9 NICKEL IONS PER MOLECULE.
 CC -1- SIMILARITY: BELONGS TO THE HTPB/HOPM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: I24513; AAA1763.1; -
DR InterPro: IPR002894; -
DR Pfam: PF01495; HyPB_Ureg; 1.
KW Nickel.
FT DOMAIN 16 54 HIS-RICH
SQ SEQUENCE 302 AA; 32737 MW; 92A4EFE58E013D01 CRC64;

Query Match 34.1%; Score 57; DB 1; Length 302;
Best Local Similarity 37.0%; Pred. No. 0.93;
Matches 10; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Oy 1 HGHEQOHGLGHGKFKLDDLEHOGH 27
DB 16 HAHSHSHAAHG--AEADSDHPAHGH 42

RESULT 14
ID HYPB_RHOCA STANDARD; PRT; 335 AA.
AC P26410;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 01-NOV-1997 (rel. 35, Last annotation update)
DE HYDROGENASE EXPRESSION/FORMATION PROTEIN HYPB.
GN HYPB OR HUPM.
OS Rhodospirillum rubrum (Rhodospirillum rubrum).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;
OC Rhodospirillum.
OX NCBI_TaxID=1061;

RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN-B10;
RX MEDLINE-93268090; PubMed=8497190;
RA Colbeau A., Richard P., Toussaint B., Caballero F.J., Elster C.,
RA Delplanck C., Smith R.L., Chabert J., Vignais P.M.,
RT "Organization of the genes necessary for hydrogenase expression in
RT Rhodospirillum rubrum mutants.";
RT hyp regulator mutants.";
RT Mol. Microbiol. 8:15-29(1993).

RN [2]
RP SEQUENCE FROM N.A.
RT MEDLINE-91177833; PubMed=2007559;
RA Xu H.W., Wall J.D.;
RT "Clustering of genes necessary for hydrogen oxidation in Rhodospirillum
RT capsulatus.";
RT J. Bacteriol. 173:2401-2405(1991).
CC -1- FUNCTION: COULD BE INVOLVED IN NICKEL BINDING AND ACCUMULATION.
CC -1- SIMILARITY: BELONGS TO THE HYPB/HUPM FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X61007; CAA43326.1; -
DR EMBL: M55089; AAA72926.1; -
DR PIR: D38532; D38532.
DR PIR: S21903; S21903.
DR InterPro: IPR002894; -
DR Pfam: PF01495; HyPB_Ureg; 1.
KW Nickel.
SQ SEQUENCE 335 AA; 35342 MW; B7276C3E1A0FD02 CRC64;

Query Match 34.1%; Score 57; DB 1; Length 335;
Best Local Similarity 40.7%; Pred. No. 1;
Matches 11; Conservative 1; Mismatches 13; Indels 2; Gaps 1;

Oy 1 HGHEQOHGLGHGKFKLDDLEHOGH 27
DB 64 HAHSHSHAAHG--AEADSDHPAHGH 88

RESULT 15
ID ZNUA_HAEIN STANDARD; PRT; 337 AA.
AC P44526;
DT 01-NOV-1995 (rel. 32, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last annotation update)
DE HIGH-AFFINITY ZINC UPTAKE SYSTEM PROTEIN ZNUA.
GN ZNUA OR H10119.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;

RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE-95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weisman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fritchman J.L., Geophagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RT Science 269:496-512(1995).
CC -1- FUNCTION: INVOLVED IN THE HIGH-AFFINITY ZINC UPTAKE TRANSPORT
CC SYSTEM (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE FIMA/PSA/SSAB/SCAA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U32698; AAC21794.1; -
DR TIGR: H10119; -
DR InterPro: IPR001987; -
DR Pfam: PF01297; Lipoprotein_4; 1.
KW Zinc; Transport.
FT DOMAIN 115 163 HIS-RICH.
SQ SEQUENCE 337 AA; 37659 MW; 3DBB45AB8F06FFCB CRC64;

Query Match 34.1%; Score 57; DB 1; Length 337;
Best Local Similarity 37.0%; Pred. No. 1;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Oy 1 HGHEQOHGLGHGKFKLDDLEHOGH 27
DB 125 HDHAKHGHKHKHDKHDDHDKHGH 151

Search completed: July 6, 2001, 09:26:38
Job time: 969 sec

AC 09VU00-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CAUP PROTEIN.
 DE CAUP OR CG10605.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amannatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Bradton R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
 RA Beeson K.Y., Bens P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chertis J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durdin K.J., Evangelista C.C., Ferraz C.C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Geldart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
 RA Jaiswal B., Jalilush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebe J.M.,
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M.C., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER HOMEOBOX DOMAINS.
 DR EMBL: AE003540; AAF49895.1; -;
 DR FLYBase: FBgn0015919; caup.
 DR InterPro: IPR001356; -;
 DR Pfam: PF00046; homeobox_1.
 DR PROSITE: PS00027; HOMEOBOX_1; 1.
 DR PROSITE: PS00071; HOMEOBOX_2; 1.
 DR SMART: SM00389; HOX_1.
 DR DNA-Binding: Homeobox; Nuclear protein.
 KW SEQUENCE 693 AA; 73667 MW; FBBB1616493F7EC9 CRC64;

	RESULT	3		
OQGTN1	OQGTN1		PRELIMINARY;	PRT; 206 AA.
AC	OQGTN1			
DT	01-MAR-2001	(TREMBLrel_16,	Created)	
DT	01-MAR-2001	(TREMBLrel_16,	Last sequence update)	
DT	01-MAR-2001	(TREMBLrel_16,	Last annotation update)	
DE	D506238.4-LIKE PROTEIN	(FRAGMENT).		
OS	Drosophila simulans	(Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachytera; Muscomorpha;			
OC	Empidoidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7240;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-NORTH CAROLINA 17;			
RA	Schmid R.J., Aguadro C.F.;			
RT	"The evolutionary analysis of 'orphans' from the Drosophila genome			
RT	identified incorrectly annotated and rapidly evolving genes."			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DDB databases.			
DR	EMBL: AF264919; AAC10258.1; -.			
FT	NON_TER	206	206	
QO	SEQUENCE	206 AA;	23011 MW;	E24B8C189BC0746D CRC64;

Query Match	42.28;	Score 70.5;	DB 5;	Length 206;
Best Local Similarity	46.78;	Pred. No. 0.029;		
Matches 14; Conservative	1;	Mismatches 12;	Indels 3;	Gaps 1

QY 1 HGEHQHGLGHC--HKFKLDDLEHQGGH 27
||| ||| : | | ||
Db 172 HGHVHGHCHGSSSHSYSLQDHHGHGH 201

RESULT	4			
Q9V3P9				
ID	Q9V3P9	PRELIMINARY;	PRT;	218 AA.
AC	Q9V3P9;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	BG:DS06238.4 PROTEIN.			
GN	BG:DS06238.4 OR CG3474.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephyridioidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-BREKLEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananidides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sultion G.G., Wortman J.E., Rendell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,			
RA	Arill J.F., Agagyanl A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Bailly R.M., Besu A.D., Bakendale J., Bayraktaroglu U., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.A., Bouck J., Brocktein P., Brotlier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Daenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K.E., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Farber C., Gabrielian A.E., Gacy N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			

RA Jallali M, Kalush F., Karpen G.H., Ke Z., Kennison J.A., Keithum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pachek J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shee B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye Y., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-Y, AND CN BW SP;
RC STRAIN-Y, AND CN BW SP;
RX MEDLINE:99403001. PubMed=14717107;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartwell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,
RA Ceiniker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RL *Drosophila melanogaster*: the *Adh* region.";
RN Genetics 153:179-219(1999).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-Y, AND CN BW SP;
RC Ceiniker S.E., Abbeyani A., Arcalata T., Baxter E., Blazej R.G.,
RA Buehnoff C., Champe N., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummalt S.R., Kaira K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomoltan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pachek J.M., Park S., Pfeiffer B., Poon L., Sequelita A.,
RA Sethi H., Smit E., Svirska R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zietan L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003643; AAF53397.1; -;
DR EMBL: AE003409; AAF44875.1; -;
DR FLYBase: FBgn0028871; BG:DS06238.4.
DR InterPro: IPR000618; -;
DR InterPro: IPR002395; -;
DR Pfam: PF00379; Insect_cuticle; 1.
DR PRINTS: PR00947; CUTICLE.
DR PRINTS: PR00334; KININOGEN.
DR PROSITE: PS00233; CUTICLE; 1.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 24371 MW; 8C920EC351E529 CRC64;

Query Match 42.2%; Score 70.5; DB 5; Length 218;
Best Local Similarity 38.9%; Pred. NO. 0.031;
Matches 14; Conservative 3; Mismatches 10; Indels 9; Gaps 1

OY 1 HIGH-----EQHGGLGHGKFKRLDDLEHOGGH 27
||| :||| ||||| | ||| :
Db 182 HGHGSSSHSYSLKQEHGHHGHSHGDHGFEGHGHY 217

RESULT 5
ID O9YWSO PRELIMINARY; PRT; 686 AA.
O9YWSO:
DT 01-MAY-2000 (TREMBlrel. 13. Created)
DT 01-MAY-2000 (TREMBlrel. 13. Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16. Last annotation update)
DE CG6632 PROTEIN.

GN CG6632.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sulton G.G., Wortman J.R., Landell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu I., Beasley E.M.,
 RA Beeson K.Y., Benson P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadenat L.B., Davies P.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mocharry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sliden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirsakis R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*."
 RT Science 287:2185-2195(2000).
 RL EMBL: AE003509; AAF48868.1; -
 DR FLYBase: FBgn0030945; CG6632.
 DR InterPro: IPR000104; -
 DR InterPro: IPR000169; -
 DR InterPro: IPR001965; -
 DR InterPro: IPR002395; -
 DR Pfam: PF00628; PHD; 1.
 DR PRINTS: PR00308; ANTIREFREEZE1.
 DR PRINTS: PR00334; KININOGEN.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN1.
 DR SMART: SM00249; PHD; 1.
 SO SEQUENCE 686 AA; 70647 MW; 17C56F19B5D2B901 CRC64.

[illegible]

ID 09JKN1 PRELIMINARY; PRT; 378 AA.
AC 09JKN1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ZINC TRANSPORTER LIKE 2.
GN ZNTL2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Zhu W., Mager S.;
RT "Cloning of new mammalian zinc transporter like genes."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF233322; AAF43423.1; -;
DR InterPro; IPR002524; -;
DR Pfam; PF01545; Cation_efflux; 1.
SQ SEQUENCE 378 AA; 41790 MW; B98AC19CA045006 CRC64;

Query Match 41.0%; Score 68.5; DB 11; Length 378;
Best Local Similarity 48.1%; Pred. No. 0.11;
Matches 13; Conservative 2; Mismatches 11; Indels 1; Gaps 1;

QY 2 GHEDQGLGHGKFKL-DDDLEHGCH 27
DB 163 GHGSHGSHGSHSLFNGALDHSCH 189

RESULT 7
Q9GTN0 PRELIMINARY; PRT; 213 AA.
AC Q9GTN0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE DS06238.4-LIKE PROTEIN (FRAGMENT).
OS Drosophila yakuba (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephydroidea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX NCBI_TaxID=7245;
RN [1]
RP SEQUENCE FROM N.A.
RA Schmid K.J., Aguadro C.F.;
RT "The evolutionary analysis of 'orphans' from the Drosophila genome
RT identifies incorrectly annotated and rapidly evolving genes."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF264920; AAG10259.1; -;
DR NON_TER 213
FT SEQUENCE 213 AA; 23731 MW; 315B8590D97886C9 CRC64;

Query Match 40.4%; Score 67.5; DB 5; Length 213;
Best Local Similarity 43.3%; Pred. No. 0.079; Indels 3; Gaps 1;
Matches 13; Conservative 2; Mismatches 12; Indels 3; Gaps 1;

QY 1 HGHEDQGLGHG--HKFKLDDLEHGCH 27
DB 179 HGHGHAHGHGSSSHSYSLKQEHGCH 208

RESULT 8
ID Q9VWX5 PRELIMINARY; PRT; 385 AA.
AC Q9VWX5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CG5936 PROTEIN.

GN CG5936.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Plankkoc C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Danlike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Modary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacliel J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodedge T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AE003508; AAF48812.1; -;
DR Flybase; FBgn0030901; CG5936.
DR InterPro; IPR002476; -;
DR Pfam; PF00001; 7cm_1; 2.
DR G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 385 AA; 43204 MW; ECAC1B1A9AFAFBE CRC64;

Query Match 39.5%; Score 66; DB 5; Length 385;
Best Local Similarity 55.0%; Pred. No. 0.24;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 HGHEDQGLGHGKFKLDDDD 20
DB 217 HGHGHAHGHGSHGLKDDDD 236

RESULT 9
ID Q9NNV9 PRELIMINARY; PRT; 198 AA.
AC Q9NNV9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)

```

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE HYPOThETICAL 23.5 KDA PROTEIN (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_Taxid=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA Spielmann T., Beck H.;
RT "Analysis of stage specific transcription in Plasmodium falciparum
RT reveals a set of genes exclusively transcribed in ring stage
RT parasites.";
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ290925; CMB92934.2;
DR InterPro: IPR002395;
DR PRINTS: PR00334; KININOGEN.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 198
SQ SEQUENCE 198 AA; 23470 MW; 4DCD4D7EE2E4BF72 CRC64;

Query Match 39.2%; Score 65.5; DB 5; Length 198;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 14; Conservative 1; Mismatches 13; Indels 7; Gaps 1;

QY 1 HGHGQHGHCCHGKFKLDDLEH-----QGSHV 28
DB 163 HAHELHDHGHGHDHGHGHDHGHGCHGV 197

RESULT 10
ID Q9W3L3 PRELIMINARY; PRT; 1937 AA.
AC Q9W3L3;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE FEMALE STERILE (1) HOMEOSTIC PROTEIN.
GN FS(1)H OR CG2252.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe C., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dletz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

```

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swircas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003442; AF46312.2;
DR Flybase; FBgn0004656; fs(1)h.
DR InterPro: IPR000104;
DR InterPro: IPR001487;
DR InterPro: IPR002173;
DR InterPro: IPR002395;
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00308; ANTI-FREZEL.
DR PRINTS; PR00503; BROMODOMAIN.
DR PRINTS; PR00334; KININOGEN.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
DR PROSITE; PS00583; PFRB_KINASES_1; UNKNOWN_1.
SQ SEQUENCE 1937 AA; 195339 MW; 1D80AA7B351F06B CRC64;

Query Match 38.9%; Score 65; DB 5; Length 1937;
Best Local Similarity 47.8%; Pred. No. 1.8;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 HGHGQHGHCCHGKFKLDDLEH 23
DB 608 HGHGHCCHGHCCHGCGSSSLKH 630

RESULT 11
ID Q9ZRC7 PRELIMINARY; PRT; 99 AA.
AC Q9ZRC7;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, last annotation update)
DE ACTINORIZAL NODULIN AGNOD-GHRP.
GN AGN84.
OS Alnus glutinosa (Alder).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fagales; Betulaceae; Alnus.
NCBI_Taxid=3517;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT NODULES;
RC TISSUE=ROOT NODULES;
RA Dobritsa S.V., Mullin B.C.;
RT "In vitro expression of actinorhizal nodulin AGNOD-GHRP and
RT demonstration of its toxicity to Escherichia coli.";
RL (in) Stacey G., Mullin B.C., Gresshoff P.M. (eds.);
RL the Biology of Plant-Microbe Interactions:
RL Proceedings of the 8th International Symposium on Molecular
RL Plant-Microbe Interactions, pp.1-1, Unknown Publisher (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT NODULES;
RC TISSUE=ROOT NODULES;
RA Twigg P.G.;
RT "Isolation of a nodule-specific cDNA encoding a putative glycine-rich
RT protein from Alnus glutinosa.";
RL Thesis (1993), The University of Tennessee, Knoxville, TN, USA.
RN [3]
SQ SEQUENCE FROM N.A.

```

```

RC TISSUE=ROOT NODULES;
RA Pawlowski K., Twigg P.G., Dobritsa S.V., Guan C., Mullin B.C.;
RT "A node-specific gene family from Alnus glutinosa encodes glycine
RT and histidine-rich proteins expressed in the early stages of
RT actinorhizal node development.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U69156; AAD00171.1; -.
DR InterPro: IPR002395; -.
DR PRINTS: PR00334; KININGEN.
DR SEQUENCE 99 AA; 10567 MW; 2ACBEAD57C070E83 CRC64;

Query Match          38.3%; Score 64; DB 10; Length 99;
Best Local Similarity 48.1%; Pred. No. 0.11;
Matches 13; Conservative 1; Mismatches 11; Indels 2; Gaps 1;

QY 1 HGHGQHGHLGHGKFKDDLEHGGH 27
DB 50 HGRHVRHGHGHVH--GNGEHGHH 74

RESULT 12
081050 PRELIMINARY; PRT; 776 AA.
AC 081050;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE T18E12.9 PROTEIN.
GN T18E12.9
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicaceae; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Rensing C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC T18E12 genomic sequence.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005313; AAC34478.1; -.
DR SEQUENCE 776 AA; 89775 MW; BE30603ACFAD14E CRC64;

Query Match          38.3%; Score 64; DB 10; Length 776;
Best Local Similarity 50.0%; Pred. No. 0.93;
Matches 15; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

QY 1 HGHGQHGHLGH--GHRFKLDDLEHGGH 28
DB 76 HGHGQHGHLGHGHSDDDDIEGIRHV 105

RESULT 13
09V5N1 PRELIMINARY; PRT; 1064 AA.
AC 09V5N1; Q9V5N2; Q24184; Q24187;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE P50 PROTEIN.
GN P50 OR CG2368.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=OVARY;
RX MEDLINE=95220671; PubMed=7705633;

```

```

RA Horowitz H., Berg C.A.;
RT "Aberrant splicing and transcription termination caused by P element
RT insertion into the intron of a Drosophila gene.";
RL Genetics 139:327-335(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC TISSUE=OVARY;
RX MEDLINE=96232300; PubMed=8674425;
RA Horowitz H., Berg C.A.;
RT "The Drosophila pipsqueak gene encodes a nuclear BRB-domain-containing
RT protein required early in oogenesis.";
RL Development 122:1859-1871(1996).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS A AND 2).
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceimiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Dugas A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dunn-Rocha S., Dunkov B.C., Dunn P.,
RA Fodor B.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann A.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegyan C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Switzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -i- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A/1 (SHOWN HERE), B AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL: U48358; AAC47153.1; -.
DR EMBL: U48402; AAC47154.1; -.
DR EMBL: AE003829; AAF58769.1; -.
DR EMBL: AE003829; AAF58770.1; -.
DR Flybase; FBgn0004399; psq.
DR InterPro; IPR002010; -.
DR InterPro; IPR002197; -.
DR Pfam; PF00651; BRB; 1.
DR PROSITE; PS50097; BRB; 1.
DR SMART; SM00225; BRB; 1.
DR Alternative splicing.
FT VARSPLIC 1 429 MISSING (IN ISOFORM B).
FT VARSPLIC 719 736 MISSING (IN ISOFORM 2).
FT CONFLICT 1020 1020 Q->QQ (IN REF. 1 AND 2).
SQ SEQUENCE 1064 AA; 114984 MW; 77420C782DEBCA5 CRC64;

```

Query Match 37.4%; Score 62.5; DB 5; Length 1064;
 Best Local Similarity 44.4%; Pred. No. 2.1;
 Matches 12; Conservative 0; Mismatches 4; Indels 11; Gaps 1;

OY 1 HGHBOQHGLGHGKFKLDDLEHOGCH 27
 DB 332 HEHEHNNHGHGH-----GGH 347

RESULT 14

O24455 PRELIMINARY; PRT; 1085 AA.
 ID Q24455; Q24456; Q24003;
 AC Q24455; Q24456; Q24003;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE PIPSOUEAK PROTEIN (BTB-V PROTEIN DOMAIN).
 GN PSQ OR CG2368.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=96134923; PubMed=8557044;
 RA Weber U., Siegel V., Mlodzik M.;
 RT "Pipsueak encodes a novel nuclear protein required downstream of
 seven-up for the development of photoreceptors R3 and R4.";
 RL EMBO J. 14:6247-6257(1995).
 RN [2]
 RP SEQUENCE OF 8-105 FROM N.A.
 RX MEDLINE=95024186; PubMed=7938017;
 RA Zolman S., Godt D., Prive G.G., Coudere J.L., Laski F.A.;
 RT "The BTB domain, found primarily in zinc finger proteins, defines an
 evolutionarily conserved family that includes several developmentally
 regulated genes in Drosophila.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10721-10722(1994).
 RN [3]
 RP ALTERNATIVE PRODUCTS; 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC EMBL; X90986; CAA62473.1; -;
 DR EMBL; X90986; CAA62474.1; -;
 DR EMBL; U14402; AAA50837.1; -;
 DR Flybase: FBgn0004399; psq.
 DR InterPro: IPR000210; -;
 DR InterPro: IPR002197; -;
 DR Pfam: PF00651; BTB; 1.
 DR PROSITE: PS50097; BTB; 1.
 DR SMART: SM00225; BTB; 1.
 KM Alternative splicing
 FT VARSPIC 428 335
 FT VARSPIC 428 335
 FT NSGLGVPKGLDFTPEPPASATPVTEHYDLSNCPSTD
 FT TRDLSTNTEPLDIDNHLAQOIHRLDQSPMHSISHHTGDS
 FT NSNLVQIKSEVIEAKHLAAOHIALS -> VSLASFTGLVS
 FT KSOQISOLARDSOISPOARMNINFLIPISLSTRIWI
 FT ACRTACMACPMDSWTSHOHHLPPDSRSTSCIPATL
 FT ARRTATYQIPEPNRSI (IN SHORT ISOFORM).
 FT MISSING (IN SHORT ISOFORM).
 FT VARSPIC 536 1084
 FT SEQUENCE 1085 AA; 117039 MW; EF32BFC73C2B737D CRC64;

Query Match 37.4%; Score 62.5; DB 5; Length 1085;
 Best Local Similarity 44.4%; Pred. No. 2.1;
 Matches 12; Conservative 0; Mismatches 4; Indels 11; Gaps 1;

OY 1 HGHBOQHGLGHGKFKLDDLEHOGCH 27
 DB 332 HEHEHNNHGHGH-----GGH 347

RESULT 15
 ID 086731 PRELIMINARY; PRT; 314 AA.
 AC 086731;

DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, last annotation update)
 DE HYPOTHETICAL 33.6 KDA PROTEIN.
 GN SC5C7.34.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_Taxid=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL031515; CAA20646.1; -;
 DR InterPro: IPR000051; -;
 KM Hypothetical protein.
 SO SEQUENCE 314 AA; 33586 MW; 7CA328BC28FF007 CRC64;

Query Match 37.1%; Score 62; DB 2; Length 314;
 Best Local Similarity 76.9%; Pred. No. 0.68;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 HGHBOQHGLGHGCH 13
 DB 18 HGHGOGHGHGHGH 30

Search completed: July 6, 2001, 09:25:54
 Job time: 990 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:09:18 ; Search time 113.68 Seconds
(without alignments)
14.932 Million cell updates/sec

Title: US-09-437-912-6

Perfect score: 177

Sequence: 1 GHKHHGHGKHKHKNGKNGKNGKNGKMT 28

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	177	100.0	28	21	AAV81997
2	177	100.0	94	21	AAV93351
3	177	100.0	179	21	AAV93353
4	177	100.0	186	21	AAV93349
5	177	100.0	255	21	AAV93342
6	177	97.2	47	21	AAV93345
7	172	97.2	62	21	AAV93348
8	172	97.2	63	16	AAV93348
9	172	97.2	83	21	AAV93347
10	172	97.2	131	16	AAV93347
11	142	80.2	41	16	AAV75180

12	142	80.2	110	16	AAV75178
13	125	70.6	20	17	AAV07625
14	98.5	55.6	19	21	AAV71879
15	98	55.4	16	21	AAV06337
16	98	55.4	16	21	AAV81999
17	98	55.4	12	21	AAV81994
18	74	41.8	12	21	AAV81995
19	72.5	41.0	177	21	AAV42194
20	72.5	41.0	180	19	AAV72390
21	69	39.0	179	21	AAV24334
22	69	39.0	179	21	AAV46905
23	68	38.4	11	21	AAV93352
24	67	37.9	1213	17	AAV06086
25	67	37.9	1213	18	AAV25029
26	66.5	37.6	85	13	AAV26414
27	66	37.3	69	16	AAV75179
28	64.5	36.4	330	21	AAV22265
29	64.5	36.4	330	21	AAV43480
30	64.5	36.4	344	21	AAV22264
31	64.5	36.4	344	21	AAV43479
32	64.5	36.4	398	21	AAV22263
33	64.5	36.4	398	21	AAV43478
34	63	35.6	351	13	AAV24393
35	62	35.0	309	21	AAV06065
36	62	35.0	309	21	AAV22955
37	62	35.0	389	21	AAV06064
38	62	35.0	389	21	AAV22954
39	62	35.0	425	21	AAV06063
40	62	35.0	453	21	AAV22953
41	62	35.0	1189	15	AAV56496
42	61.5	34.7	1085	17	AAV95607
43	61	34.5	149	16	AAV76478
44	59.5	33.6	294	21	AAV28680
45	59.5	33.6	374	21	AAV28679

ALIGNMENTS

RESULT 1	
AAV81997	
ID	AAV81997 standard; peptide; 28 AA.
AC	AAV81997;
DT	16-OCT-2000 (first entry)
DE	Human high molecular weight kininogen domain 5 fragment #6.
XX	Human; high molecular weight kininogen; HK;
KW	two-chain high molecular weight kininogen; HKa;
KW	angiogenesis inhibition; tumour; cancer; ocular disorder;
KW	rheumatoid arthritis; endothelial cell apoptosis.
OS	Homo sapiens.
PN	WO200027866-A1.
PD	18-MAY-2000.
XX	
PF	05-NOV-1999; 99WO-US26419.
XX	
PR	10-NOV-1998; 98US-0107833.
XX	
PA	(UTEM) UNIV TEMPLE.
XX	(MCCR/) MCCRAE R. K.
PI	McCrae RK;
XX	
DR	WPI; 2000-376483/32.
XX	
PT	A pharmaceutical composition used to inhibit angiogenesis, inhibit endothelial cell proliferation, and induce endothelial cell apoptosis

[illegible]

```

xx OS Synthetic.
xx OS Homo sapiens.
xx PN WO200027415-A2.
xx PD 18-MAY-2000.
xx PE 09-NOV-1999; 99WO-US26377.
xx PR 10-NOV-1998; 98US-0107844.
xx PA (UTEM ) UNIV TEMPLE.
xx PA (DUPO ) DUPONT PHARM CO.
xx PA (COLM/) COLMAN W. R.
xx PA (MOUN/) MOUNSA A. S.
xx PT Colman WR, Mounsa AS.
xx DR WPI, 2000-376306/32.
xx PT Method for inhibiting endothelial cell proliferation, using compound
xx PT that inhibit endothelial cell migration
xx PS Claim 6; Page 37; 41pp; English.
xx PS The present sequence represents an analogue of the light chain of human
xx CC high molecular weight kininogen. High molecular weight kininogen is a
xx CC 120 kDa glycoprotein which binds with high affinity to endothelial cells,
xx CC where it is cleaved by plasma kallikrein into heavy and light chains.
xx CC Analogues of high molecular weight kininogen are used in the method
xx CC of the invention. The specification describes a method of inhibiting
xx CC endothelial cell proliferation. The method comprises contacting
xx CC endothelial cells with a compound containing high molecular weight
xx CC kininogen analogues. The method and the compounds can be used for
xx CC inhibiting endothelial cell proliferation. The compounds can also be
xx CC used for inhibiting angiogenesis. The compounds can also be used to
xx CC inhibit migration of endothelial cells to vitronectin.
xx SQ Sequence 62 AA:

Query Match 97.2%; Score 172; DB 21; Length 62;
Best Local Similarity 100.0%; Pred. No. 1,7e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 GHKHKHGHHGKHKKNKKGNKNGWK 27
   |||||
Db 36 ghkhkghghgkhkhkngkngkhngwk 62

RESULT 8
AAR75186
AAR75186 standard; peptide; 63 AA.
AAR75186;
AAR75186;
05-DEC-1995 (first entry)
Partial peptide of human HMW kininogen fragment 2.
high molecular weight; kininogen; fragment; 1,2; 1,2; partial;
KW wound treating agent; bovine; growth promotion; fibroblast.
Homo sapiens.
xx OS Homo sapiens.
xx PN JP07082172-A.
xx PD 28-MAR-1995.
xx PF 17-SEP-1993; 93JP-0230616.
xx PR 17-SEP-1993; 93JP-0230616.

```

XX	PA	(FARH) HOECHST JAPAN KK.
XX	DR	WPI: 1995-158909/21.
XX	PT	A wound treating agent contg. a partial peptide of kininogen -
XX	PT	have growth promotion activity of fibroblasts.
XX	PS	Claim 8; Page 8; 8pp; Japanese.
XX	CC	AAE75186 is a partial peptide corresponding to human kininogen
XX	CC	fragment 1, amino acids 458-520. Partial peptides of bovine and
XX	CC	human kininogen fragments 1,2, 1 and 2, are used in wound treating
XX	CC	agent compns. and act as the active component. The fragments are
XX	CC	useful in wound treating because they have growth promotion activity
XX	CC	on fibroblasts.
XX	SQ	Sequence 63 AA;
OY		Query Match 97.2%; Score 172; DB 16; Length 63;
		Best Local Similarity 100.0%; Pred. No. 1.7e-14;
		Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Db	1	GHRKHGHHGKHKRNGKNGKNGKNGWK 27
	37	ghkhnhgngkhkngkngkngkngwk 63
RESULT 9		
AA993347.		
ID	AA93347	standard; peptide; 83 AA.
XX	AC	AA93347;
XX	DT	04-SEP-2000 (first entry)
XX	DE	Light chain of human high molecular weight kininogen analogue.
XX	KW	Human; high molecular weight kininogen; glycoprotein; endothelial cell;
KW	KW	plasma kallikrein; heavy chain; light chain; analogue; angiogenesis;
KW	KW	endothelial cell proliferation; endothelial cell migration; vitronectin.
OS	OS	Synthetic.
OS	OS	Homo sapiens.
XX	PN	WO200027415-A2.
XX	PD	18-MAY-2000.
XX	PF	09-NOV-1999; 99WO-0526377.
XX	PR	10-NOV-1998; 98US-0107844.
XX	PA	(UTEM) UNIV TEMPLE.
PA	PA	(DUPO) DUPONT PHARM CO.
PA	PA	(COLM/) COLMAN W R.
PA	PA	(MOSA/) MOUSA A S.
XX	PI	Colman WR, Mousa AS;
XX	DR	WPI: 2000-376306/32.
XX	PT	Method for inhibiting endothelial cell proliferation, using compound
XX	PT	that inhibit endothelial cell migration -
XX	PS	Claim 5; Page 37; 41pp; English.
XX	CC	The present sequence represents an analogue of the light chain of human
XX	CC	high molecular weight kininogen. High molecular weight kininogen is a
XX	CC	120 kDa glycoprotein which binds with high affinity to endothelial cells
XX	CC	where it is cleaved by plasma kallikrein into heavy and light chains.
XX	CC	Analogues of high molecular weight kininogen are used in the method

CC	of the invention. The specification describes a method of inhibiting
CC	endothelial cell proliferation. The method comprises contacting
CC	endothelial cells with a compound containing high molecular weight
CC	kininogen analogues. The method and the compounds can be used for
CC	inhibiting endothelial cell proliferation. The compounds can also be
CC	used for inhibiting angiogenesis. The compounds can also be used to
CC	inhibit migration of endothelial cells to vitronectin.
XX	
SO	Sequence 83 AA:
QY	Query Match 97.2%; Score 172; DB 21; Length 83;
	Best Local Similarity 100.0%; Pred. No. 2.3e-14;
Matches	27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GHKRHGHGCHKHKKKKGNKNGWK 27
Db	57 gkhkhghghgkhkhkkgkngkhgwk 83
RESULT 10	
AA75181	
ID	AA75181 standard; peptide: 131 AA.
XX	
AC	AA75181;
XX	
DT	05-DEC-1995 (first entry)
XX	
DE	Partial peptide of human HMW kininogen fragment 1.2.
XX	
KW	high molecular weight; kininogen; fragment; 1.2; 1; 2; partial;
KW	wound treating agent; human; growth promotion; fibroblast.
XX	
OS	Homo sapiens.
XX	
PN	JP07082172-A.
XX	
PD	28-MAR-1995.
XX	
PF	17-SEP-1993; 93JP-0230616.
XX	
PR	17-SEP-1993; 93JP-0230616.
XX	
PA	(FARH) HOECHST JAPAN KK.
XX	
DR	WPI: 1995-158909/21.
XX	
PT	A wound treating agent contg. a partial peptide of kininogen -
PT	have growth promotion activity of fibroblasts.
XX	
PS	Claim 7; Page 7; 8pp; Japanese.
XX	
XX	
CC	AA75181 is a partial peptide corresponding to human kininogen
CC	fragment 1.2, amino acids 390-520. Partial peptides of bovine and
CC	human kininogen fragments 1.2, 1 and 2, are used in wound treating
CC	agent compns. and act as the active component. The fragments are
CC	useful in wound treating because they have growth promotion activity
CC	on fibroblasts.
XX	
XX	
SO	Sequence 131 AA:
QY	Query Match 97.2%; Score 172; DB 16; Length 131;
	Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches	27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GHKRHGHGCHKHKKKKGNKNGWK 27
Db	105 gkhkhghghgkhkhkkgkngkhgwk 131
RESULT 11	
AA75180	

ID	AA075180 standard; peptide; 41 AA.
XX	
AC	AA075180;
XX	
DT	05-DEC-1995 (first entry)
XX	
DE	Partial peptide of HMW kininogen fragment 2.
XX	
KW	high molecular weight; kininogen; fragment; 1.2; 1; 2; partial; wound treating agent; bovine; growth promotion; fibroblast.
XX	
OS	Bos taurus.
XX	
PN	JP07082172-A.
XX	
PD	28-MAR-1995.
XX	
Pf	17-SEP-1993; 93JP-0230616.
XX	
PR	17-SEP-1993; 93JP-0230616.
XX	
PA	(FARH) HOECHST JAPAN KK.
XX	
DR	WPI; 1995-156809/21.
XX	
PT	A wound treating agent contg. a partial peptide of kininogen - have growth promotion activity of fibroblasts.
XX	
PS	Claim 6; Page 7; 8pp; Japanese.
XX	
CC	AA075179 is a partial peptide corresponding to bovine kininogen fragment 2, amino acids 456-496. Partial peptides of bovine and human kininogen fragments 1.2, 1 and 2, are used in wound treating agent compsns. and act as the active component. The fragments are useful in wound treating because they have growth promotion activity on fibroblasts.
XX	
SQ	Sequence 41 AA;
Query Match	80.2%; Score 142; DB 16; Length 41;
Best Local Similarity	81.5%; Pred. No. 5,2e-11;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;	
OY	1 GHKHKHGHCGRKKKGKKNKNGWK 27 : Db 15 ghkhghngghkhnkgknngkydwr 41
RESULT 12	
AA075178	
ID	AA075178 standard; peptide; 110 AA.
XX	
AC	AA075178;
XX	
DT	05-DEC-1995 (first entry)
XX	
DE	Partial peptide of HMW kininogen fragment 1.2.
XX	
KW	high molecular weight; kininogen; fragment; 1.2; 1; 2; partial; wound treating agent; bovine; growth promotion; fibroblast.
XX	
OS	Bos taurus.
XX	
Key	Location/Qualifiers
FH	Misc-difference 12 /label= Pro, Thr
FT	Misc-difference 15 /label= Val or Leu
FT	Misc-difference 69 /label= Lys or Arg
XX	
JN	JP07082172-A.

[illegible]

XX Claim 3; Page 2; 14pp; Japanese.
 PS
 CC The present peptide, and its claimed fragments, are derived from
 CC residues 402-498 of the human high polymer quininogen L-chain. They
 CC are useful in cell adhesion, cancer metastasis or platelet
 CC aggregation inhibitors, and in wound, inflammatory disease,
 CC arteriosclerosis or glomerular nephritis treating agents. The
 CC present peptide was synthesised using a solid phase method, and
 CC purified using a YMC-DOS-120A-S15/13 column.
 XX
 SQ Sequence 20 AA;
 Query Match 70.6%; Score 125; DB 17; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 4 HKHGHGKHKKNGKNGKH 23
 1 hkhghghkknkngkngkh 20
 DB
 RESULT 14
 AAY71879
 ID AAY71879 standard; peptide; 19 AA.
 XX
 AC AAY71879;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human HKH20 peptide derived from domain 5 of H-Kininogen (479-498 aa).
 XX
 KW Human; heparin binding protein; HBP; antiinflammatory; cardiovascular;
 KW immunosuppressive; vasotropic; prevention; treatment; bradykinin;
 KW apelin; H-Kininogen; HK; systemic inflammatory response syndrome;
 KW pre-kallikrein; ischemia reperfusion; anaphylaxis; allograft rejection;
 KW adult respiratory distress syndrome; HKH20 peptide.
 XX
 OS Homo sapiens.
 XX
 PN WO200066151-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 28-APR-2000; 2000WO-DK00213.
 XX
 PR 29-APR-1999; 99GS-0132748.
 PR 06-MAY-1999; 99DK-0000613.
 PR 01-OCT-1999; 99DK-0001402.
 PR 01-OCT-1999; 99US-0157384.
 XX
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Flodgaard HJ, Lindbom L, Bjorn S;
 XX
 DR WPI; 2000-687445/67.
 XX
 PT Treating systemic inflammatory response syndrome, ischaemia reperfusion,
 PT anaphylaxis and allograft rejection by modulating release of bradykinin
 XX
 PS Example 2; Page 39; 75pp; English.
 XX
 CC The patent discloses a method for preventing or treating a disorder
 CC resulting from the release of bradykinin in a mammal which produces
 CC a heparin-binding protein (HBP) that binds to a HBP antagonist. This
 CC method involves administration of a mammalian HBP antagonist (especially
 CC apelinogen) and/or monoclonal antibodies that bind to prekallikrein-
 CC H-kininogen complexes in the HBP, to decrease the release of bradykinin
 CC in the mammal. The antagonists of HBP (e.g. apelinogen) decrease the
 CC permeability of the endothelial cells and are used to prevent or treat
 CC disorders resulting from the release of bradykinin such as systemic

CC inflammatory response syndrome, ischaemia reperfusion, anaphylaxis
 CC and/or allograft rejection. They are also used to treat adult
 CC respiratory distress syndrome.
 CC
 CC The present sequence is HKH20 peptide which is derived from the
 CC domain 5 of human H-Kininogen (HK) protein (479-498 residues).
 CC HKH20 treatment of endothelial cells inhibits or prevents the HBP-
 CC induced increase in permeability of the endothelial cells.
 XX
 SQ Sequence 19 AA;
 Query Match 55.6%; Score 98.5; DB 21; Length 19;
 Best Local Similarity 90.0%; Pred. No. 4.8e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 OY 4 HKHGHGKHKKNGKNGKH 23
 1 hkhghghk-knkngkngkh 19
 DB
 RESULT 15
 AAB06337
 ID AAB06337 standard; Protein; 16 AA.
 XX
 AC AAB06337;
 XX
 DT 16-OCT-2000 (first entry)
 XX
 DE Human two-chain high molecular weight kininogen domain 5 fragment #9.
 XX
 KW Human; high molecular weight kininogen; HK;
 KW two-chain high molecular weight kininogen; HKa;
 KW angiotensin inhibition; tumour; cancer; ocular disorder;
 KW rheumatoid arthritis; endothelial cell apoptosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200027866-A1.
 XX
 PD 18-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US26419.
 XX
 PR 10-NOV-1998; 98US-0107833.
 XX
 PA (UTEM) UNIV TEMPLE.
 PA (MCCR/) MCCRAE R K.
 PI MCCRAE RK;
 XX
 DR WPI; 2000-376483/32.
 XX
 PT A pharmaceutical composition used to inhibit angiogenesis, inhibit
 PT endothelial cell proliferation, and induce endothelial cell apoptosis
 XX
 PS Claim 15; Page 29; 52pp; English.
 XX
 CC The present sequence is derived from human two-chain high molecular
 CC weight kininogen (Hka) domain 5. Hka is product of high molecular
 CC weight kininogen (HK) cleavage by plasma kallikrein. HK is a 120 kd
 CC glycoprotein which binds with high affinity to endothelial cells. Hka
 CC or a synthetic compound comprising the present sequence may be used in
 CC a pharmaceutical composition for inhibiting angiogenesis. Angiogenesis
 CC occurs in a number of disease states, such as tumour formation and
 CC expansion, and certain ocular disorders. It can also occur in a
 CC rheumatoid joint, hastening joint destruction by allowing an influx
 CC of leukocytes. The composition may inhibit angiogenesis by inhibiting
 CC endothelial cell proliferation or by inducing endothelial cell
 CC apoptosis. Peptides used in the composition may be recombinant peptides,
 CC natural peptides, or synthetic peptides. They may also be chemically
 CC synthesised, using, for example, solid phase synthesis methods.
 XX

SO Sequence 16 AA;

Query Match 55.4%; Score 98; DB 21; Length 16;

Best Local Similarity 100.0%; Pred. No. 4.7e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 HKKGGKKRKHNGWKT 28
 Db 1 hknkgkkngkhngwkt 16

Search completed: July 6, 2001, 09:09:18
 Job time: 124 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:10:22 ; Search time 56.74 Seconds

(without alignments)
9.941 Million cell updates/sec

Title: US-09-437-912-6

Perfect score: 177
Sequence: 1 GHKHHGHGHGKHKNGKNGKNGKNGWT 28

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	37.9	1213	1	US-08-188-582-20
2	67	37.9	1213	1	US-08-646-715-20
3	61.5	34.7	1085	1	US-08-431-080-28
4	61.5	34.7	1085	2	US-08-938-534-28
5	59	33.3	515	4	US-08-942-012B-32
6	59	33.3	1199	4	US-09-208-742-2
7	56.5	31.9	313	3	US-08-686-528A-3
8	56.5	31.9	313	4	US-09-456-287-3
9	56.5	31.9	337	3	US-08-686-528A-2
10	56.5	31.9	337	4	US-09-456-287-2
11	56.5	31.9	376	1	US-08-594-031-100
12	56.5	31.9	376	1	US-08-594-031-102
13	55	31.1	60	1	US-08-235-457-1
14	55	31.1	60	2	US-09-115-032-1
15	55	31.1	60	5	PCR-US95-05772-1
16	54.5	30.8	1261	4	US-09-208-742-4
17	54	30.5	834	2	US-08-861-464-4
18	54	30.5	834	2	US-08-396-001-4
19	54	30.5	834	4	US-09-323-433A-4
20	53	29.9	28	1	US-08-152-488-2
21	53	29.9	28	1	US-08-303-025-2
22	53	29.9	28	1	US-08-677-304-2
23	53	29.9	28	2	US-08-436-703B-7
24	53	29.9	617	1	US-08-137-614A-26
25	53	29.9	637	3	US-08-072-064-1
26	53	29.9	637	3	US-08-072-064-4
27	53	29.9	637	3	US-08-072-064-6

28	53	29.9	637	3	US-08-072-064-8	Sequence 8, Appl1
29	53	29.9	637	5	PCR-US92-08558-1	Sequence 1, Appl1
30	51.5	29.1	53	2	US-08-651-818A-19	Sequence 19, Appl1
31	51.5	29.1	54	2	US-08-651-818A-23	Sequence 23, Appl1
32	51.5	29.1	452	1	US-08-434-702-6	Sequence 6, Appl1
33	51.5	29.1	452	1	US-08-271-883-6	Sequence 6, Appl1
34	51.5	29.1	1032	4	US-09-115-954-8	Sequence 8, Appl1
35	51.5	29.1	1044	4	US-09-115-954-2	Sequence 2, Appl1
36	51	28.8	400	5	PCR-US95-16472-2	Sequence 2, Appl1
37	51	28.8	402	3	US-08-602-809-2	Sequence 2, Appl1
38	51	28.8	651	1	US-08-431-080-24	Sequence 24, Appl1
39	51	28.8	651	1	US-08-938-534-24	Sequence 24, Appl1
40	50	28.2	113	2	US-08-918-727-7	Sequence 7, Appl1
41	50	28.2	113	3	US-09-205-680A-7	Sequence 7, Appl1
42	50	28.2	345	2	US-08-758-621-14	Sequence 14, Appl1
43	50	28.2	345	4	US-09-107-858-14	Sequence 14, Appl1
44	50	28.2	575	2	US-09-032-315-8	Sequence 8, Appl1
45	50	28.2	575	2	US-08-993-318A-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-188-582-20
; Sequence 20, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comal, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TATS AND METHODS OF USE
; CORRESPONDENCE ADDRESS: 36
; ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188, 582
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1213 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-188-582-20
Query Match 37.9%; Score 67; DB 1; Length 1213;

Best Local Similarity 57.9%; Pred. No. 1.7;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKHKGHGKHKNGKN 20
||||| | | | | :
DB 1160 HKHKGHRHRSKDKERKD 1178

RESULT 2
US-08-646-715-20

; Sequence 20, Application US/08646715
; Patent No. 5637686

; GENERAL INFORMATION:

; APPLICANT: Tjian, Robert

; APPLICANT: Comai, Lucio

; APPLICANT: Dynlacht, Brian D.

; APPLICANT: Hoey, Timothy

; APPLICANT: Ruppert, Siegfried

; APPLICANT: Tanese, Naoko

; APPLICANT: Weinzierl, Robert O.J.

; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/646,715

; FILING DATE: 09-MAY-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/188,582

; FILING DATE: 28-JAN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman, Richard A.

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1213 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-646-715-20

Query Match 37.9%; Score 67; DB 1; Length 1213;

Best Local Similarity 57.9%; Pred. No. 1.7;

Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKHKGHGKHKNGKN 20
||||| | | | | :
DB 1160 HKHKGHRHRSKDKERKD 1178

RESULT 3

US-08-431-080-28

; Sequence 28, Application US/08431080

; Patent No. 5698686

; GENERAL INFORMATION:

; APPLICANT: Gottschling, Daniel E.

; APPLICANT: Singer, Miriam S.

; TITLE OF INVENTION: Telomerase Compositions and Methods

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TEXAS

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/431,080

; FILING DATE: Concurrently Herewith

; CLASSIFICATION: 514

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: SN 08/326,781

; FILING DATE: October 20, 1994

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, David L.

; REGISTRATION NUMBER: 32,165

; REFERENCE/DOCKET NUMBER: ARCD:155/PAR

; TELEPHONE: (512) 418-3000

; TELEFAX: (713) 789-2679

; TELEX: 79-0924

; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1085 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-431-080-28

Query Match 34.7%; Score 61.5; DB 1; Length 1085;

Best Local Similarity 48.3%; Pred. No. 6.9;

Matches 14; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

QY 1 GKR-----HKHKGHGKHKNGKN 24
||| | | | | :
DB 512 GKRKGRHRSKSHLEKNGSNLIKSN 540

RESULT 4
US-08-938-534-28

; Sequence 28, Application US/08938534

; Patent No. 5916752

; GENERAL INFORMATION:

; APPLICANT: Gottschling, Daniel E.

; APPLICANT: Singer, Miriam S.

; TITLE OF INVENTION: Telomerase Compositions and Methods

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TEXAS

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:


```

; APPLICATION NUMBER: US/08/938,534
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,080
; FILING DATE:
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-938-534-28

Query Match          34.7%; Score 61.5; DB 2; Length 1085;
Best Local Similarity 48.3%; Pred. No. 6.9;
Matches 14; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

QY      1 GHK-----HKHGHGKHKHKKKNGKHN 24
      ||| ||| ||| ||| ||| ||| ||| |||
DB      512 GHKSKKGRHKSCKSHLEHKNKSNLTKSN 540

RESULT 5
US-08-942-012B-32
; Sequence 32, Application US/08942012B
; Patent No. 6235278
; GENERAL INFORMATION:
; APPLICANT: Miller, Lois K.
; APPLICANT: Lu, Albert
; APPLICANT: Dierks, Peter
; APPLICANT: Black, Bruce
; TITLE OF INVENTION: Biological Insect Control Agents Expressing
; TITLE OF INVENTION: Insect-Specific Toxin Genes, Methods and Compositions
; FILE REFERENCE: 28-96A
; CURRENT APPLICATION NUMBER: US/08/942,012B
; CURRENT FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: 08/729,606
; PRIOR FILING DATE: 2000-10-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 32
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Spodoptera littoralis nuclear polyhedrosis virus
;
US-08-942-012B-32

Query Match          33.3%; Score 59; DB 4; Length 515;
Best Local Similarity 57.1%; Pred. No. 6.6;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 HKHKGHGKHKHKN 15
      ||||| ||| |||
DB      58 HNNHGHRRHREN 71

RESULT 6
US-09-208-742-2
; Sequence 2, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:
```

```

; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: CIP150/hTAF1150 is Necessary for Cell
; TITLE OF INVENTION: Cycle Progression
; FILE REFERENCE: 1453.002
; CURRENT APPLICATION NUMBER: US/09/208,742
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1199
; TYPE: PRT
; ORGANISM: human
;
US-09-208-742-2

Query Match          33.3%; Score 59; DB 4; Length 1199;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 11; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      2 HKHKGHGKHKHKKKNGKHN 23
      ||| ||| ||| ||| ||| ||| ||| |||
DB      1146 HHHHHHEHKKKKKKHKKHKKH 1167

RESULT 7
US-08-686-528A-3
; Sequence 3, Application US/08686528A
; Patent No. 6054134
; GENERAL INFORMATION:
; APPLICANT: LINGWOOD, Clifford A.
; TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,528A
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Rea, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 032609-001
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-686-528A-3

Query Match          31.9%; Score 56.5; DB 3; Length 313;
Best Local Similarity 43.5%; Pred. No. 8;
Matches 10; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY      2 HKHKGHGKHKHKKKNGKHN 24
      ||||| ||| ||| ||| ||| ||| ||| |||
DB      113 HKHHDHHDH-DHKHKKHDEHNN 134

RESULT 8
US-09-456-287-3
```


US-08-594-031-100
; Sequence 100, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; US-08-594-031-100

Query Match 31.9%; Score 56.5; DB 1; Length 376;
Best Local Similarity 47.6%; Pred. No. 9.5;
Matches 10; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

OY 7 GHGKGKKNKGGKKNKNGWK 27
I:| | | | | :| | | | |
Db 104 GNGTGSNN--VDGKHGWR 121

RESULT 12
US-08-594-031-102
; Sequence 102, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,031
FILING DATE: 30-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-594-031-102

Query Match 31.9%; Score 56.5; DB 1; Length 376;
Best Local Similarity 47.6%; Pred. No. 9.5;
Matches 10; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

OY 7 GHGKGKKNKGGKKNKNGWK 27
I:| | | | | :| | | | |
Db 104 GNGTGSNN--VDGKHGWR 121

RESULT 13
US-08-255-457-1
; Sequence 1, Application US/08255457
; Patent No. 578040
; GENERAL INFORMATION:
; APPLICANT: Plaut, Andrew G.
; APPLICANT: Gilbert-Rothstein, Joanne V.
; APPLICANT: Wright, Andrew
; TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/255,457
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul C.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00398/090001

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-255-457-1

Query Match 31.1%; Score 55; DB 1; Length 60;
Best Local Similarity 37.5%; Pred. No. 2.4;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 GHKHKHGCHGKKNKGNKGN 24
DB 10 GHHNNHNTHTNNHYGSEHHNNH 33

RESULT 14
US-09-115-032-1
Sequence 1, Application US/09115032
Patent No. 5972348
GENERAL INFORMATION:
APPLICANT: Plaut, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,032
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/255,457
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-115-032-1

Query Match 31.1%; Score 55; DB 2; Length 60;
Best Local Similarity 37.5%; Pred. No. 2.4;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 GHKHKHGCHGKKNKGNKGN 24
DB 10 GHHNNHNTHTNNHYGSEHHNNH 33

RESULT 15
PCT-US95-05772-1
Sequence 1, Application PC/TUS9505772
GENERAL INFORMATION:
APPLICANT: Plaut, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL
TITLE OF INVENTION: BINDING PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05772
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-05772-1

Query Match 31.1%; Score 55; DB 5; Length 60;
Best Local Similarity 37.5%; Pred. No. 2.4;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 GHKHKHGCHGKKNKGNKGN 24
DB 10 GHHNNHNTHTNNHYGSEHHNNH 33

Search completed: July 6, 2001, 09:10:23
Job time: 189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2001, 09:18:01 ; Search time 73.59 Seconds

(without alignments)
28,983 Million cell updates/sec

Title: US-09-437-912-6

Sequence: 1 GHKHHGHGKHKHCKNGKNGKNGKMT 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR,68:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	177	100.0	644	1	KGHUHL	kininogen, HMW pre
2	147	83.1	619	1	KGBOH2	kininogen, HMW II
3	147	83.1	621	1	KGBOH1	kininogen, HMW I P
4	88	49.7	264	2	C25486	K-kininogen, HMW P
5	88	49.7	639	2	A25486	kininogen, HMW I P
6	80.5	45.5	290	2	C27115	K-kininogen, LMW P
7	80.5	45.5	315	2	A27115	major acute phase
8	74.5	41.1	199	2	T48099	hypothetical prote
9	73	41.0	507	2	S54303	zinc transport pro
10	72.5	41.0	177	2	S65780	glycine/proline-r1
11	71.5	40.4	173	2	A51469	glycine/proline-r1
12	71	40.1	2038	2	A43742	female sterile hom
13	69.5	39.3	503	2	S54302	zinc transporter 2
14	69.5	39.3	515	2	T23089	hypothetical prote
15	69	39.0	98	2	S08137	gene 2C protein -
16	69	39.0	179	2	A85217	hypothetical prote
17	69	39.0	251	2	T34168	hypothetical prote
18	67.5	38.1	670	2	F36791	hypothetical prote
19	67	37.9	213	2	S04491	dermal gland prote
20	67	37.9	1213	2	A54063	TkR-binding prote
21	66.5	37.6	85	2	A45969	hemolymph antifun
22	66	37.3	125	2	T49356	hypothetical prote
23	66	37.3	229	2	T27840	hypothetical prote
24	66	37.3	336	1	S75947	hypothetical prote
25	66	37.3	420	1	DC8CD	diminopimelate de
26	66	37.3	420	2	B85936	diminopimelate de
27	66	37.3	1891	2	T13594	hypothetical prote
28	65	36.7	102	2	T07078	cold stress protei
29	65	36.7	220	2	A44805	eggshell protein p

30	64.5	36.4	398	2	T02681	probable zinc tran
31	64.5	36.4	549	2	T15506	hypothetical prote
32	64	36.2	207	2	T08109	oleosin-like prote
33	64	36.2	375	2	T08134	oleosin-like prote
34	63.5	35.9	110	2	T07618	cold stress protei
35	63.5	35.9	196	2	D85999	hypothetical prote
36	63.5	35.9	196	2	A49987	probable fkbp-type
37	63.5	35.9	226	2	T27843	hypothetical prote
38	63	35.6	208	2	T08132	oleosin-like prote
39	63	35.6	286	2	S07193	chorion protein s3
40	63	35.6	351	1	KGZOH	histidine-rich gly
41	63	35.6	735	2	T45059	hypothetical prote
42	62.5	35.3	306	2	S08607	chorion protein s3
43	62.5	35.3	980	1	S45444	BNM1 protein-bind
44	62	35.0	189	2	C81428	peptidyl-prolyl ci
45	62	35.0	389	2	B96635	hypothetical prote

ALIGNMENTS

RESULT 1

KGHUHL

kininogen, HMW precursor [validated] - human

N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen

N:Contents: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular

C:Species: Homo sapiens (man)

C:date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000

C:Accession: A01279; A25276; S32422; A24871; A27899; A27699; A31905; A34030;

R:Okubo, I.; Kurachi, K.; Takasawa, T.; Shikawa, H.; Sasaki, M.

Biochemistry 23, 5691-5697, 1984

A:Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide

A:Reference number: A90490; MUID:8512621

A:Accession: A01279

A:Molecule type: mRNA

A:Residues: 1-389 <OHK>

A:Cross-references: GB:K02566; NID:9177889

R:Takagaki, Y.; Kitamura, N.; Nakanishi, S.

J. Biol. Chem. 260, 8601-8609, 1985

A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and I

A:Reference number: A92544; MUID:85234582

A:Accession: A25276

A:Molecule type: mRNA

A:Residues: 1-592, 'T', 594-644 <TK>

A:Cross-references: GB:M11437; NID:9186751; PIDN:AAB59550.1; PID:9386852

R:Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.

FEBS Lett. 321, 93-97, 1993

A:Title: Cloning, expression and characterization of human kininogen domain 3.

A:Reference number: S32422; MUID:93223854

A:Accession: S32422

A:Molecule type: mRNA

A:Residues: 'ANSM', 253-377 <NUE>

A:Note: differences are due to known cloning artifacts

R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foerisch, B.; Muller-Esterl, W.

Eur. J. Biochem. 152, 307-314, 1985

A:Title: The amino acid sequence of the light chain of human high-molecular-mass kini

A:Reference number: A91153; MUID:86030270

A:Accession: A91153

A:Molecule type: protein

A:Residues: 379-644 <LOT>

A:Note: the bradykinin sequence preceding the light chain sequence was not determined

R:Kellermann, J.; Lottspeich, F.; Henschen, A.; Muller-Esterl, W.

Eur. J. Biochem. 154, 471-478, 1986

A:Title: Completion of the primary structure of human high-molecular-mass kininogen.

A:Reference number: A24871; MUID:86108361

A:Accession: A24871

A:Molecule type: protein

A:Residues: 'Z', 20-380 <KEL>

R:Kellermann, J.; Lottspeich, F.; Henschen, A.; Muller-Esterl, W.

In Kinins IV, Greenbaum, U.M., and Margolis, H.S., ed., pp.85-89, Plenum Press, New

A:Title: Amino acid sequence of the light chain of human high molecular mass kininoge

A:Reference number: A27899

A:Accession: A27899

J. Biochem. 79, 1201-1222, 1976
A>Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino acid sequence has been determined from complementary DNA.
A:Reference number: A91941; MUID:76260155
A:Accession: A91941
A:Molecule type: protein
A:Residues: 387-455 <HAN>
A>Note: 398-Pro, 401-Val, and 455-Lys were also found
R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A>Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Amide bond cleavage sites.
A:Reference number: A91938; MUID:75170265
A:Accession: A91938
A:Molecule type: protein
A:Residues: 456-496 <HA2>
R:Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga, S.
J. Biol. Chem. 252, 2768-2779, 1987
A>Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of carboxylate groups, and the amino acid sequence of the light chain.
A:Reference number: A92627; MUID:87137530
A:Accession: B29359
A:Molecule type: protein
A:Residues: 'Z', '20-104', 'E', '106-256', 'XX', '257-376 <SUE>
R:Riettsch, F.; Kellerman, J.; Henschel, A.; Foerster, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A>Title: The amino acid sequence of the light chain of human high-molecular-mass kininogen.
A:Reference number: A91153; MUID:86030270
A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R:Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A>Title: Disulfide bonds in bovine HMW kininogens.
A:Reference number: A94300
A:Contents: annotation; disulfide bonds
A>Note: article in Japanese
C:Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as shown by cDNA sequencing.
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the blood coagulation cascade.
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is important for its biological activity.
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, 1-oxo-octadecanamide, is present in the kininogen prior to the release of bradykinin.
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; duplicated; duplication; family; kininogen; signal sequence; status predicted <SIG>
F:1-18/Domains: signal sequence #status predicted <SIG>
F:19-619/Product: HMW kininogen II #status predicted <MAV>
F:19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F:19-130/Domains: cystatin homology <CY1>
F:141-252/Domains: cystatin homology <CY2>
F:261-372/Domains: cystatin homology <CY3>
F:337-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F:387-386/Product: bradykinin (kallidin I) #status experimental <BKY>
F:387-619/Product: HMW kininogen II light chain #status experimental <LCH>
F:418-488/Region: glycine/histidine/lysine-rich
F:19/Moderified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental <MOD>
F:127-589, 82-93, 106-125, 141-144, 205-217, 228-247, 261-264, 325-337, 348-367/Disulfide bonds:
F:87, 166, 169, 204, 280/Binding site: carbonylate (Asn) (covalent) #status absent
F:147/Binding site: carbonylate (Thr) (covalent) #status experimental
F:136/Binding site: carbonylate (Thr) (covalent) (partial) #status experimental
F:19/Binding site: carbonylate (Asn) (covalent) (partial) #status experimental
F:376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
F:380/Moderified site: 4-hydroxyproline (Pro) #status predicted
F:386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:396, 400, 404, 510/Binding site: carbonylate (Ser) (covalent) #status experimental
F:397, 398, 518, 522, 534, 546, 551, 568/Binding site: carbonylate (Thr) (covalent) #status experimental
F:496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match	83.1%	Score 147,	DB 1;	Length 619;
Best Local Similarity	82.1%	Pred. No. 7.8e-10;		
Matches	23;	Conservative 1;	Mismatches 4;	Indels 0;
			Gaps 0;	
Oy	1	GHKHKHGHHGKHKKNKGNKNGMT	28	
Db	470	GHGKHKHGHHGKHKKNKGNKNGHIDWRT	497	
RESULT	3			

KGB001
 N:kininogen, HMW I precursor - bovine
 N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
 N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
 N:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Nov-1993 #sequence:revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A01281; A91923; A91938; A29555
 R:Kitamura, N.; Takasaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
 Nature 305, 545-549, 1983
 A:Title: A single gene for bovine high molecular weight and low molecular weight kininogens
 A:Reference number: A93317; MUID:84014106
 A:Accession: A01281
 A:Molecule type: mRNA
 A:Residues: 1-621 <KIT>
 A:Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA4735.1; PID:g492
 R:Kato, H.; Nagasawa, S.; Suzuki, T.
 J. Biochem. 67, 313-323, 1970
 A:Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds a
 A:Reference number: A91923; MUID:70180420
 A:Accession: A91923
 A:Molecule type: protein
 A:Residues: 378-393 <KAT>
 R:Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
 J. Biochem. 77, 55-68, 1975
 A:Title: Studies on the primary structure of bovine high-molecular-weight kininogen.
 A:Reference number: A91938; MUID:75170265
 A:Accession: A91938
 A:Molecule type: protein
 A:Residues: 458-498 <HAN>
 R:Suwayoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwana
 J. Biol. Chem. 262, 2768-2779, 1987
 A:Title: Bovine high molecular weight kininogen. The amino acid sequence, positions o
 A:Reference number: A92627; MUID:87137530
 A:Accession: A29559
 A:Molecule type: protein
 A:Residues: 2, 20-123, '1, 125-127, '1, 129-378 <SUD>
 R:Lotzspeich, F.; Kellermann, J.; Henschel, A.; Foerstsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A:Title: The amino acid sequence of the light chain of human high-molecular-mass kinin
 A:Reference number: A91153; MUID:86030270
 A:Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
 R:Suwayoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
 Seikagaku 56, 808, 1984
 A:Title: Disulfide bonds in bovine HMW kininogens.
 A:Reference number: A94300
 A:Contents: annotation; disulfide bonds
 A:Note: article in Japanese
 C:Comment: The HMW kininogen precursor is produced from the LMW form
 C:Comment: kininogen is a cysteine proteinase inhibitor, takes part in initiation of
 C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is i
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; d
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-621/Product: HMW prokininogen I #status predicted <MAT>
 F:19-378/Product: HMW kininogen I heavy chain #status experimental <HCH>
 F:19-130/Domain: cystatin homology <CYI>
 F:141-252/Domain: cystatin homology <CY2>
 F:263-374/Domain: cystatin homology <CY3>
 F:379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <RBDY>
 F:389-388/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:389-621/Product: HMW kininogen I light chain #status experimental <LCH>
 F:417-468/Region: glycine/histidine/lysine-rich
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:27-591, 82-93, 106-125, 141-144, 205-217, 228-247, 263-266, 327-339, 350-369/Disulfide bond
 F:87, 168, 169, 204/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
 F:197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:378-379/Cleavage site: Met-Lys (Kallikrein) #status experimental
 F:382/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:388-389/Cleavage site: Arg-Ser (Kallikrein) #status experimental
 F:398-399/Binding site: carbohydrate (Ser) (covalent) #status experimental
 F:399, 406, 512/Binding site: carbohydrate (Ser) (covalent) #status experimental

F.417-021/Product: www.kallikrein.com Lygic acid #status experimental

F.417-488/Region: glycine/histidine/lysine-rich #status experimental

F.191/Modified site: pyrrolidone carboxylic acid (Gln) (in zwitter form) #status experi
F.27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bond
F.87,168,169,204/Binding site: carboxylate (Asn) (covalent) #status experimental
F.116/Binding site: carboxylate (Thr) (covalent) (partial) #status experimental
F.197/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental
F.378-379/Cleavage site: Met-Lys (Kallikrein) #status experimental
F.382/Modified site: 4-Hydroxyproline (Pro) #status predicted
F.388-389/Cleavage site: Arg-Ser (Kallikrein) #status experimental
F.398,406,512/Binding site: carboxylate (Ser) (covalent) #status experimental

RT "Disulfide bonds in bovine HMW kininogens." ;

Seikagaku 56:508-808(1984).

-i- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THROMBOPLASTIN, (2) HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN AND PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE CONTRACTION, (4B) INDUCTION OF HYPERTENSION, (4C) NATEPHRESIS AND DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION, INDIRECTLY VIA ENDOTHELIUM-DEIVED RELAXING FACTOR ACTION); (5) LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD CLOTTING.

-i- SUBCELLULAR LOCATION: SECRETED.

-i- ALTERNATIVE PRODUCTS: 2 ISOFORMS: HMW (SHOWN HERE) AND LMW; ARE PRODUCED BY ALTERNATIVE SPLICING.

-i- TISSUE SPECIFICITY: PLASMA.

-i- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.

-i- SIMILARITY: CONTAINS 3 CYSTEATIN-LIKE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

```

RL      SeiKakaku 56:808-808(1984).
CC      -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES. (2)
CC      HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC      HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT
CC      FACTOR XII: (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN
CC      INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC      BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC      PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC      CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NARCIPTURES AND
CC      DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC      MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC      PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC      OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT IS
CC      A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC      INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC      LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW
CC      KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC      CLOTTING.
CC      -1- SUBCELLULAR LOCATION: SECRETED.
CC      -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC      PRODUCED BY ALTERNATIVE SPLICING.
CC      -1- TISSUE SPECIFICITY: PLASMA.
CC      -1- PMW: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN
CC      -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC      -----
CC      THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLAB
CC      BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTU
CC      THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS O
CC      USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN ACCO
CC      MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMM
CC      ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isb-sib.ch/ann
CC      OR SEND AN EMAIL TO license@isb-sib.ch).
CC      -----
DR      EMBL; K02566; AAA35497.1; -.
DR      EMBL; M11437; AAB59550.1; -.
DR      EMBL; M11438; AAB59550.1; JOINED.
DR      EMBL; M11521; AAB59550.1; JOINED.
DR      EMBL; M11522; AAB59550.1; JOINED.
DR      EMBL; M11523; AAB59550.1; JOINED.
DR      EMBL; M11524; AAB59550.1; JOINED.
DR      EMBL; M11525; AAB59550.1; JOINED.
DR      EMBL; M11526; AAB59550.1; JOINED.
DR      EMBL; M11527; AAB59550.1; JOINED.
DR      EMBL; M11528; AAB59550.1; JOINED.
DR      EMBL; M11528; AAB59550.1; JOINED.
DR      EMBL; M11437; AAB59551.1; -.
DR      EMBL; M11438; AAB59551.1; JOINED.
DR      EMBL; M11521; AAB59551.1; JOINED.
DR      EMBL; M11522; AAB59551.1; JOINED.
DR      EMBL; M11523; AAB59551.1; JOINED.
DR      EMBL; M11524; AAB59551.1; JOINED.
DR      EMBL; M11525; AAB59551.1; JOINED.
DR      EMBL; M11526; AAB59551.1; JOINED.
DR      EMBL; M11527; AAB59551.1; JOINED.
DR      EMBL; M11528; AAB59551.1; JOINED.
DR      PIR; A01279; KGH0H1.
DR      PIR; A25276; A25276.
DR      PIR; A01280; KGH0L1.
DR      PIR; B25276; B25276.
DR      PIR; S02482; S02482.
DR      SWISS-2DPAGE; P01043; HUMAN.
DR      MIM; 228960; -.
DR      InterPro; IPR000010; -.
DR      InterPro; IPR002395; -.
DR      Pfam; PF00031; cystatin.3.
DR      PRINTS; PF00334; KININOGEN.
DR      PROSITE; PS00287; CYSTATIN.2.
KW      Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator
KW      Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW      Alternative splicing.
FT      SIGNAL          1      18
FT      CHAIN           19      644      KININOGEN.
FT      CHAIN           19      380      KININOGEN HEAVY CHAIN.
FT      PEPTIDE         381      389      BRADYKININ.

```

FT	CHAIN	390	644	KININOGEN LIGHT CHAIN.
FT	DOMAIN	19	136	CYSTATIN-LIKE 1.
FT	DOMAIN	137	258	CYSTATIN-LIKE 2.
FT	DOMAIN	259	380	CYSTATIN-LIKE 3.
FT	DOMAIN	420	510	HIS-RICH (ASSOCIATED WITH CLOTTING ACTIVITY).
FT	REPEAT	420	449	
FT	REPEAT	450	479	
FT	REPEAT	480	510	
FT	MOD_RES	19	19	
FT	DISULEFID	28	614	
FT	DISULEFID	83	94	
FT	DISULEFID	107	126	
FT	DISULEFID	142	145	
FT	DISULEFID	206	218	
FT	DISULEFID	229	248	
FT	DISULEFID	264	267	
FT	DISULEFID	328	340	
FT	DISULEFID	351	370	
FT	CARBOHYD	48	48	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	205	205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	294	294	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	401	401	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	533	533	
FT	CARBOHYD	542	542	
FT	CARBOHYD	546	546	
FT	CARBOHYD	557	557	
FT	CARBOHYD	571	571	
FT	CARBOHYD	577	577	
FT	CARBOHYD	593	593	
FT	CARBOHYD	628	628	
FT	VARSPPLIC	402	427	
FT	VARSPPLIC	428	644	VSPPTMAPADDERDSGKDGQHTR -> SHLRSCYEYKGR
FT	CONFLIC	593	593	PKKAGAEPAEREVEV (IN ISOFORM LMW).
FT	SEQUENCE	644 AA; 71945 MW; 31323BCBAAFBFBVE CRC64;		MISSING (IN ISOFORM LMW).
FT	SEQUENCE			T -> I (IN REF. 1).

Query Match	100.0%;	Score 177;	DB 1;	Length 644;
Best local Similarity	100.0%;	Pred. No. 1e-14;		
Matches	28;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	GHRKHGHGHGKHKNKGKKNKGKKNKGWKT	28	
Db	494	GHRKHGHGHGKHKNKGKKNKGKKNKGWKT	521	
RESULT	2			
KNH2_BOVIN		STANDARD:	PRT;	619 AA.
ID	KNH2_BOVIN			
AC	P01045;			
DT	21-JUL-1986 (rel. 01, created)			
DT	21-JUL-1986 (rel. 01, last sequence update)			
DT	01-OCT-2000 (rel. 40, last annotation update)			
DE	KININOGEN, HMW II PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
NCBI_TaxID	9913;			
1)				
SEQUENCE FROM N.A.				
RP	MEDLINE-84014106; PubMed-6571699;			
RA	Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;			
RT	"A single gene for bovine high molecular weight and low molecular weight kininogens."			
RL	Nature 305:545-549 (1983).			
12)				
RP	SEQUENCE OF 19-376.			
RP	MEDLINE-87137530; PubMed-3546295;			
RA	Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.;			

RN Miyata T., Iwanaga S.;
RT "Bovine high molecular weight kininogen. The amino acid sequence,
RT positions of carbohydrate chains and disulfide bridges in the heavy
RT chain portion."
RL J. Biol. Chem. 262:2768-2779(1987).
RN [3]
RP SEQUENCE OF 376-391.
RX MEDLINE=70180420; PubMed=4986212;
RA Kato H., Nagasawa S., Suzuki T.;
RT "Studies on the structure of bovine kininogen: cleavages of disulfide
RT bonds and of methionyl bonds in kininogen-II";
RL J. Biochem. 67:313-323(1970).
RN [4]
RP SEQUENCE OF 387-455.
RX MEDLINE=76260155; PubMed=956151;
RA Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT "Primary structure of bovine plasma high-molecular-weight kininogen.
RT The amino acid sequence of a glycopeptide portion (fragment 1)
RT following the C-terminus of the bradykinin moiety";
RL J. Biochem. 79:1201-1222(1976).
RN [5]
RP SEQUENCE OF 456-496.
RX MEDLINE=75170265; PubMed=1169237;
RA Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT "Studies on the primary structure of bovine high-molecular-weight
RT peptide' released by plasma kallikrein.";
RL J. Biochem. 77:55-68(1975).
CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THROMBOPLASTIN-INDUCED
CC FIBRINOLYSIS. (2) KININOGENS ARE INHIBITORS OF THROMBIN-INDUCED
CC FIBRINOLYSIS. (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND
CC PLASMIN-INDUCED AGGREGATION OF THROMBOCYTES. (4) THE ACTIVE
CC PEPTIDE BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS
CC A VARIETY OF PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH
CC MUSCLE CONTRACTION. (4B) INDUCTION OF HYPOTENSION. (4C)
CC NARROWESIS AND DIURESIS. (4D) DECREASE IN BLOOD GLUCOSE LEVEL.
CC (4E) IT IS A MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE
CC IN VASCULAR PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3)
CC RELEASE OF OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS),
CC (4F) IT HAS A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ
CC ACTION, INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR
CC ACTION).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC TO RESIDUE 398.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- PM: BRAVUKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC-----
DR EMBL, V01492; CAA24736.1; -;
DR EMBL, V01492; CAA24737.1; ALT_SEQ.
DR PIR, A01282; KGBOH2.
DR PIR, B29559; B29559.
DR HSSP, P04129; IART.
DR InterPro, IPRO00010; -;
DR InterPro, IPRO02395; -;
DR Pfam, PF00031; cystatin_3.
DR PRINTS, PR00334; KININOGEN.
DR PROSITE, PS00287; CYSTATIN_2.
KW Glycoprotein. Plasma. Duplication. Vasodilator. Alternative splicing;
KW Thiol protease inhibitor; Bradykinin. Blood coagulation; Signal;
KW Inflammatory response.

FT	SIGNAL	1	18		KININOGEN, HMM II.
FT	CHAIN	19	619		HEAVY CHAIN.
FT	CHAIN	19	376		BADRYKININ.
FT	PEPTIDE	387	386		LIGHT CHAIN.
FT	CHAIN	378	619		CYSTATIN-LIKE 1.
FT	DOMAIN	19	135		CYSTATIN-LIKE 2.
FT	DOMAIN	136	256		CYSTATIN-LIKE 3.
FT	DOMAIN	257	376		PYROLIDONE CARBOXYLIC ACID.
FT	MOD_RES	19	19		N-LINKED (GLCNAC. . .).
FT	CARBOHYD	87	87		PARTIAL.
FT	CARBOHYD	136	136		OR 169.
FT	CARBOHYD	168	168		N-LINKED (GLCNAC. . .) (PARTIAL).
FT	CARBOHYD	197	197		N-LINKED (GLCNAC. . .).
FT	CARBOHYD	204	204		N-LINKED (GLCNAC. . .).
FT	CARBOHYD	280	280		INTERCHAIN.
FT	CARBOHYD	400	400		
FT	DISULFID	27	589		
FT	DISULFID	82	93		
FT	DISULFID	106	125		
FT	DISULFID	141	144		
FT	DISULFID	205	217		
FT	DISULFID	228	247		
FT	DISULFID	261	264		
FT	DISULFID	325	337		
FT	DISULFID	348	367		T -> P.
FT	VARIANT	398	398		L -> V.
FT	VARIANT	401	401		H -> K.
FT	VARIANT	454	454		
SO	SEQUENCE	619 AA:	68710 MM:	F04320A8B0E0EDA CRC64;	
Query Match					
Best Local Similarity		83.1%;	Score 147;	DB 1;	Length 619;
Matches		23;	Conservative	1;	Mismatches
				4;	Indels
					Gaps
Oy	1 GHKHGHGKHKKKGKNKNGHMKT	28			
Db	470 GHGKHGHGKHKKKGKNKNGHMDWT	497			
RESULT 3					
KNH1_BOVIN					
ID	KNH1_BOVIN	STANDARD;	PRT;	621 AA.	
AC	P01044;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	01-JUN-1994 (Rel. 29, Last annotation update)				
DE	KININOGEN, HMM I PRECURSOR (THIOL PROTEINASE INHIBITOR) [CONTAINS: BRADYKININ].				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;				
CC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=84014106; PubMed=6571699;				
RA	Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakamishi S.;				
RT	"A single gene for bovine high molecular weight and low molecular				
RT	weight kininogens.";				
RL	Nature 305:545-549(1983).				
RN	[2]				
RP	SEQUENCE OF 19-378.				
RX	MEDLINE=87137530; PubMed=3546295;				
RA	Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,				
RA	Miyata T., Iwanaga S.;				
RT	"Bovine high molecular weight kininogen. The amino acid sequence,				
RT	positions of carbohydrate chains and disulfide bridges in the heavy				
RT	chain portion.";				
RL	J. Biol. Chem. 262:2768-2779(1987).				
RN	[3]				
RP	SEQUENCE OF 378-393.				
RX	MEDLINE=70180420; PubMed=4986212;				

DR InterPro: IPR002395; -
 DR InterPro: IPR003243; -
 DR Pfam: PF00031; cystatin. 3.
 DR PRINTS: PR00334; KININOGEN.
 DR PROSITE: PS00287; CYSTATIN. 1.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 661
 FT CHAIN 19 379
 FT CHAIN 380 388
 FT CHAIN 389 661
 FT DOMAIN 19 135
 FT DOMAIN 136 257
 FT DOMAIN 258 379
 FT DOMAIN 439 524
 FT DOMAIN 28 631
 FT DISULFID 83 94
 FT DISULFID 107 125
 FT DISULFID 141 144
 FT DISULFID 205 217
 FT DISULFID 228 247
 FT DISULFID 263 266
 FT DISULFID 327 339
 FT DISULFID 350 369
 FT CARBOHYD 82 82
 FT CARBOHYD 168 168
 FT CARBOHYD 204 204
 FT CARBOHYD 242 242
 FT VARSPLIC 401 432
 FT FT
 FT VARSPLIC 433 661
 FT SEQUENCE 661 AA: 73102 MW: 774460258D587966 CRC64:
 Query Match 52.5%; Score 93; DB 1; Length 661;
 Best Local Similarity 57.1%; Pred No. 0.0002;
 Matches 16; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
 Qy 1 GHKHKHGHGKHKRNGKNGKNGKNGKNGK 28
 DB 514 GHGHHGHGKHKTKDKNSVKQTQRT 541
 RESULT 5
 KNG_RAT STANDARD; PRT; 639 AA.
 AC P08934; P08933;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE KININOGEN PRECURSOR [CONTAINS: BRADYKININ].
 GN KNG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 OX [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RP MEDLINE=87137443; PubMed=3029068;
 RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
 RT "Differing expression patterns and evolution of the rat kininogen
 RT gene family."; J. Biol. Chem. 262:2190-2198(1987).
 RL [2]
 RP SEQUENCE FROM N.A. (LMW ISOFORM).
 RP MEDLINE=86008264; PubMed=2413018;
 RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
 RT "Primary structures of the mRNAs encoding the rat precursors for
 RT bradykinin and T-kinin. Structural relationship of kininogens with
 RT major acute phase protein and alpha 1-cysteine proteinase

RT Inhibitor.";
 RT J. Biol. Chem. 260:12054-12059(1985).
 RL [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RP STRAIN-BUFFALO;
 RX MEDLINE=87250580; PubMed=2439509;
 RA Fung W.-P., Schreiber G.;
 RT "Structure and expression of the genes for major acute phase alpha 1-
 RT protein (thioalbumin) and kininogen in the rat."; J.
 RT J. Biol. Chem. 262:9298-9308(1987).
 RN [4]
 RP SEQUENCE OF 1-41 FROM N.A.
 RP STRAIN-WISTAR; TISSUE=Liver;
 RX MEDLINE=87137465; PubMed=3818598;
 RA Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
 RT "Differing utilization of homologous transcription initiation sites
 RT of rat k and T kininogen genes under inflammation condition."; J.
 RT J. Biol. Chem. 262:2345-2351(1987).
 RL
 CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN- AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -1- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
 CC HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
 CC T-II.
 CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: L29428; AAA41486.1; -
 CC EMBL: M11884; AAA41487.1; -
 CC EMBL: M14369; AAA41484.1; -
 CC EMBL: M14369; AAA41485.1; ALT_SEQ.
 CC EMBL: M16455; AAA41482.1; -
 CC PIR: A25486; A25486.
 CC PIR: A28055; A28055.
 CC InterPro: IPR000010; -
 CC InterPro: IPR002395; -
 CC Pfam: PF00031; cystatin. 3.
 CC PRINTS: PR00334; KININOGEN.
 CC PROSITE: PS00287; CYSTATIN. 2.
 KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
 KW Alternative splicing; Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 639
 FT CHAIN 19 380
 FT PEPTIDE 381 389
 FT CHAIN 390 639
 FT KININOGEN.
 FT KININOGEN HEAVY CHAIN.
 FT BRADYKININ.
 FT KININOGEN LIGHT CHAIN.

```

FT DOMAIN 19 136 CYSTATIN-LIKE 1.
FT DOMAIN 137 258 CYSTATIN-LIKE 2.
FT DOMAIN 259 380 CYSTATIN-LIKE 3.
FT DOMAIN 439 514 HIS-RICH.
FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 107 126 BY SIMILARITY.
FT DISULFID 142 145 BY SIMILARITY.
FT DISULFID 206 218 BY SIMILARITY.
FT DISULFID 229 248 BY SIMILARITY.
FT DISULFID 264 267 BY SIMILARITY.
FT DISULFID 328 340 BY SIMILARITY.
FT DISULFID 351 370 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 402 433 VSPSYIARVOEERDPGEGPIGHGHLAKQ -> RLNS
CEYGRLLKAGAPAPAPAPAPAPASTVP (IN ISOFORM
LMN).
FT VARSPLIC 434 639 MISSING (IN ISOFORM LMN).
FT CONFLICT 61 61 E -> K (IN REF. 2).
SQ SEQUENCE 639 AA; 70933 MW; D3172DE94PF56AF5 CRC64;

```

Query Match 49.7%; Score 88; DB 1; Length 639;
 Best Local Similarity 65.2%; Pred. No. 0.0008;
 Matches 15; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

OY 1 GHKHKHGHGKHKKNGKNGK 23
DB 492 GHGHHGHGKDKHTNKDKNNYKH 514

```

```

RESULT 6
ZNTL_RAT 6
ID ZNTL_RAT STANDARD; PRT; 507 AA.
AC 062720;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ZINC TRANSPORTER 1 (ZNT-1).
GN SLC30A1 OR ZNT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=95188868; PubMed=7882967;
RA Palmer R.D., Findley S.D.;
RT "Cloning and functional characterization of a mammalian zinc
RT transporter that confers resistance to zinc."
RT EMBO J. 14:639-649(1995).
RL [2]
RP INDUCTION BY ZINC.
RP TISSUE-Intestine;
RX MEDLINE=98226729; PubMed=9560190;
RX McMahon R.O., Cousins R.J.;
RX "Regulation of the zinc transporter Znt-1 by dietary zinc."
RX Proc. Natl. Acad. Sci. U.S.A. 95:4841-4846(1998).
RL -1- FUNCTION: MAY BE INVOLVED IN ZINC TRANSPORT OUT OF THE CELL.
CC -1- SUBUNIT: MULTIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- LOCALIZED ON THE PLASMA MEMBRANE (PROBABLE). LOCALIZED ON THE
CC BASOLATERAL SURFACE OF THE ENTEROCYTES.
CC TISSUE SPECIFICITY: WIDELY EXPRESSED. THE PROTEIN IS DETECTED IN
CC DUODENUM AND JEJUNUM BUT NOT IN ILEUM AND COLON.
CC -1- INDUCTION: SLIGHTLY BY ZINC IN THE INTESTINE, BUT NOT THE LIVER.
CC -1- SIMILARITY: BELONGS TO THE SLC30A FAMILY OF TRANSPORTERS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U17133; AAA79234.1; -
DR InterPro: IPR002524; -
DR Pfam: PF01545; Cation_efflux_1
KW Zinc; Transport; Transmembrane; Multigene family; Repeat.
FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 11 31 POTENTIAL.
FT DOMAIN 32 35 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 36 56 POTENTIAL.
FT DOMAIN 57 78 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 79 99 POTENTIAL.
FT DOMAIN 100 113 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 114 134 POTENTIAL.
FT TRANSSEM 135 247 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 248 268 POTENTIAL.
FT TRANSSEM 269 307 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 308 328 POTENTIAL.
FT TRANSSEM 329 507 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 145 156 6 X 2 AA APPROXIMATE REPEATS OF H-G.
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 507 AA; 55142 MW; 9F9770017C2455FC CRC64;

```

Query Match 41.2%; Score 73; DB 1; Length 507;
 Best Local Similarity 57.7%; Pred. No. 0.044;
 Matches 15; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

```

OY 1 GHKHKHGHGKHKKNG-KKNGKNG 25
DB 145 GHGHHGHG-HLAKGARKKGRAGG 169

```

```

RESULT 7
FSH_DROME 7
ID FSH_DROME STANDARD; PRT; 2038 AA.
AC P13709; P13710;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE FEMALE STERILE HOMEOTIC PROTEIN (FRAGILE-CHORION MEMBRANE PROTEIN).
GN FS(1)H OR FSH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89276730; PubMed=2567251;
RX Haynes S.R., Mozer B.A., Bhatia-Dey N., David I.B.;
RX "The Drosophila fsh locus, a maternal effect homeotic gene, encodes
RX apparent membrane proteins."
RX Dev. Biol. 134:246-257(1989).
RL -1- FUNCTION: REQUIRED MATERNAALLY FOR PROPER EXPRESSION OF OTHER
CC HOMEOTIC GENES INVOLVED IN PATTERN FORMATION, SUCH AS UBX.
CC -1- SIMILARITY: HIGH. NO HUMAN RING3 PROTEIN.
CC -1- SIMILARITY: CONTAINS 2 BROMODOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33221; AAA8540.1; -
CC DR EMBL: M33222; AAA8541.1; ALT_TERM.
CC DR EMBL: M15762; AAA70424.1; -
CC DR EMBL: M15763; AAA70423.1; -
CC DR EMBL: M15764; AAA70422.1; -
CC PIR: A43742; A43742.
CC HSP: P04002; 1WFA.
CC FLYbase: Fp9n0004656; fs(1)h.
CC InterPro: IPR001487; -.
CC Pfam: PF00439; bromodomai; 2.
CC PRINTS: PR00503; BROMODOMAIN.
CC PROSITE: PS00633; BROMODOMAIN_1; 2.
CC PROSITE: PS50044; BROMODOMAIN_2; 2.
CC KW Developmental protein; Bromodomain; Transmembrane; Repeat.
CC FT DOMAIN 51 123 BROMODOMAIN 1.
CC FT DOMAIN 495 567 BROMODOMAIN 2.
CC FT DOMAIN 945 1106 FT DOMAIN.
CC FT TRANSMEM 330 350 POTENTIAL.
CC FT TRANSMEM 451 471 POTENTIAL.
CC FT TRANSMEM 750 770 POTENTIAL.
CC FT TRANSMEM 790 810 POTENTIAL.
CC FT TRANSMEM 816 830 POTENTIAL.
CC FT TRANSMEM 874 894 POTENTIAL.
CC FT TRANSMEM 1731 1751 POTENTIAL.
CC FT TRANSMEM 1939 1959 POTENTIAL.
CC FT VARIANT 909 909 G -> A.
CC FT VARIANT 1022 1022 H -> RRPY.
CC SQ SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match
Best Local Similarity 40.1%; Score 71; DB 1; Length 2038;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 GHKHKHGHGKHKRKNKNGKNGW 26
Db 597 GHKHKHGHGKHKRKNKNGKNGW 622

RESULT 8
KEY BRARE STANDARD; PRT; 352 AA.
AC 09PUB8:
DT 01-OCT-2000 (rel. 40, Created)
DT 01-OCT-2000 (rel. 40, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE HISTIDINE-RICH MEMBRANE PROTEIN KE4 HOMOLOG (FRAGMENT).
GN HKE4.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Suelmann H., Klein J.;
RT Identification of a homolog of the human HKE4 gene in zebrafish.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE KE4/CATSUP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF196345; AAF05821.1; -
CC KW Transmembrane; Glycoprotein.

```

```

FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 128 148 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT DOMAIN 24 105 HIS-RICH.
FT DOMAIN 177 217 HIS-RICH.
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 352 352
SQ SEQUENCE 352 AA; 37922 MW; C8C8C60F6D2BA8A6 CRC64;

Query Match
Best Local Similarity 39.3%; Score 69.5; DB 1; Length 352;
Matches 14; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 1 GHKHKHGHGKHKRKNKNGKNGW 24
Db 94 GHKHKHGHGKHKRKNKNGKNGW 120

RESULT 9
ZNT1_MOUSE STANDARD; PRT; 503 AA.
AC 060738;
DT 30-MAY-2000 (rel. 39, Created)
DT 30-MAY-2000 (rel. 39, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE ZINC TRANSPORTER 1 (ZNT-1).
GN SLC30A1 OR ZNT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
RC MEDLINE=95188868; PubMed=7882967;
RA Palmer R.D., Findley S.D.;
RT Cloning and functional characterization of a mammalian zinc
RT transporter that confers resistance to zinc.;
RL EMBO J. 14:639-649(1995).
CC -1- FUNCTION: MAY BE INVOLVED IN ZINC TRANSPORT OUT OF THE CELL.
CC -1- LEPTALITY OF KNOCKOUT EARLY IN GESTATION SUGGESTS A ROLE OF THE
CC PROTEIN IN FETAL ZINC ACQUISITION AND RETENTION.
CC -1- SUBUNIT: MULTIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC LOCALIZED ON THE PLASMA MEMBRANE (PROBABLE).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE SLC30A FAMILY OF TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U17132; AAA79233.1; -
CC DR MGD; MGI:1345281; SLC30A1.
CC DR InterPro: IPR002524; -.
CC Pfam: PF01545; Cation_efflux; 1.
CC ZINC: Transport; Transmembrane; Multigene family; Repeat.
CC FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 11 31 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 32 35 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 57 78 POTENTIAL.
CC FT TRANSMEM 79 99 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 100 113 POTENTIAL.
CC FT TRANSMEM 114 134 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 135 243 POTENTIAL.
CC FT DOMAIN 135 243 CYTOPLASMIC (POTENTIAL).

```



```

RT encode homeoproteins that control proneural and vein-forming genes.";
RL Cell 85:95-110(1996).
CC -1- FUNCTION: CONTROLS PRONEURAL AND VEIN FORMING GENES. POSITIVE
CC TRANSCRIPTIONAL CONTROLLER OF AC-SC (ACHAETE-SCUTE). MAY ACT AS AN
CC ACTIVATOR THAT INTERACTS WITH THE TRANSCRIPTIONAL COMPLEX
CC ASSEMBLED ON THE AC AND SC PROMOTERS AND PARTICIPATES IN
CC TRANSCRIPTION INITIATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE TALE/IRO FAMILY OF HOMEOBOX PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X95178; CA64485.1; -.
CC HSSP; P02833; ISAN.
CC FLYBASE; FBgn0015919; caup.
CC InterPro: IPR001356; -.
CC Pfam; PF00046; homeobox_1.
CC PROSITE; PS00027; HOMEOBOX_1; 1.
CC PROSITE; PS00071; HOMEOBOX_2; 1.
CC Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
CC Developmental protein.
CC DNA_BIND 226 288 HOMEOBOX (TALE-TYPE).
CC DOMAIN 300 303 POLY-ASP.
CC FT DOMAIN 405 418 POLY-GLN.
CC FT DOMAIN 501 516 POLY-GLN.
CC FT DOMAIN 517 528 POLY-HIS.
CC FT DOMAIN 565 572 POLY-SER.
CC FT DOMAIN 613 624 POLY-SER.
CC SEQUENCE 693 AA; 73749 MW; 8E0D6D43C9CDC619 CRC64;
-----
Query Match 39.0%; Score 69; DB 1; Length 693;
Best local similarity 48.0%; Pred. No. 0.18;
Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;
-----
Qy 1 GHKHKHGHGKHKHKKNGKNGK 25
Db 649 GHGSHGHGHHGHGHLGHLG 673
-----
RESULT 13
ID VGS0_HSV11 STANDARD; PRT; 670 AA.
AC 000130;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL GENE 50 PROTEIN.
GN 50.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.
OC unclassified Herpesviridae.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AU8URN 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M75136; AA88153.1; -.
CC PIR; F36791; F36791.
CC DR PIR; F36791; F36791.
CC KW Hypothetical protein; Repeat.
CC REPEAT 143 158
CC FT REPEAT 171 186
CC FT REPEAT 200 214
CC FT REPEAT 215 233
CC FT REPEAT 224 252
CC FT REPEAT 253 268
CC FT REPEAT 279 293
CC FT REPEAT 294 309
CC FT REPEAT 320 334
CC FT REPEAT 335 349
CC FT REPEAT 362 376
CC FT REPEAT 391 391
CC FT REPEAT 392 406
CC FT REPEAT 407 421
CC FT REPEAT 422 436
CC FT REPEAT 437 452
CC FT REPEAT 464 477
CC FT REPEAT 478 493
CC FT REPEAT 504 517
CC FT REPEAT 518 531
CC FT REPEAT 532 545
CC FT REPEAT 546 559
CC FT REPEAT 560 573
CC FT REPEAT 574 587
CC FT REPEAT 588 601
CC FT REPEAT 602 615
CC FT REPEAT 616 629
CC SEQUENCE 670 AA; 64174 MW; 2B64A781C519E8B4 CRC64;
-----
Query Match 38.1%; Score 67.5; DB 1; Length 670;
Best local similarity 57.1%; Pred. No. 0.27;
Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
-----
Qy 1 GHKHKHGHGKHKHKKNGKNGK 21
Db 642 GHGHHGHGHHG-HGGRGPPGG 661
-----
RESULT 14
ID SKGR_XENLA STANDARD; PRT; 213 AA.
AC P13673;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE SKIN GRANULE PROTEIN PRECURSOR.
DE Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=8928999; PubMed=2737290;
RA Berger H., Kell G.;
RT "The constituents of storage granules in the dermal glands of Xenopus
RT laevis. Structure of a basic polypeptide deduced from cloned cDNA.";
RL FEBS Lett. 249:293-296(1989).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```


AC 027920;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE PC4-1.
 GN PC4-1.
 OS Bradyzia hygida.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Sciaroidea;
 OC Sciaridae; Bradyzia.
 OX NCBI_TaxID=35572;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SALIVARY GLAND;
 RX MEDLINE=95393845; PubMed=7664619;
 RA Monesi N., Fernandez M.A., Fontes A.M., Basso L.R., Nakanishi Y.,
 Baron B., Buttin G., Peco-Larson M.L.,
 RT "Molecular characterization of an 18 kb segment of DNA puf C4 of
 Bradyzia hygida (Diptera, sciaridae)."
 RL Chromosoma 103:715-724(1995).
 DR EMBL; U13883; AAA83554.1; -
 DR EMBL; U13892; AAA83555.1; -
 SQ SEQUENCE 450 AA; 47185 MW; 1F0633CE9B7F964C CRC64;

Query Match 44.1%; Score 78; DB 5; Length 450;
 Best Local Similarity 55.6%; Pred. No. 0.0062;
 Matches 15; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

OY 1 GHRKHGHGHRKNGKNGKNGKNGK 27
 Db 76 GHRKHGHGHRG-HGHGHNHNGGRK 100

RESULT 3
 O9VMSO PRELIMINARY; PRT; 686 AA.
 AC 09VMSO;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CG6632 PROTEIN.
 GN CG6632.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Abmayyan A., An H.-J., Andrews-Flannock C., Baldwin D.,
 Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Morkulov G., Mlshina N.V., Mobbary C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003509; AAF48868.1; -
 DR FlyBase; FBgn0030945; CG6632.
 DR InterPro; IPR000104; -
 DR InterPro; IPR000169; -
 DR InterPro; IPR001965; -
 DR InterPro; IPR002395; -
 DR Pfam; PF00628; PHD; 1.
 DR PRINTS; PR00308; ANTIFREEZE1.
 DR PRINTS; PR00334; KININOGEN.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 DR SMART; SM00249; PHD; 1.
 SQ SEQUENCE 686 AA; 70647 MW; 17C56F19B5D28901 CRC64;

Query Match 43.5%; Score 77; DB 5; Length 686;
 Best Local Similarity 57.1%; Pred. No. 0.013;
 Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 GHRKHGHGHRKNGKNGKNGKNG 21
 Db 471 GHRKHGHGHRGSHSSGHRG 491

RESULT 4
 O9ZRC7 PRELIMINARY; PRT; 99 AA.
 AC 09ZRC7;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE ACTINORITZAL NODULIN AGNOD-GHRP.
 GN AGNT84.
 OS Alnus glutinosa (Alder).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Dobritsa S.V., Mullin B.C.;
 RT "In vitro expression of actinorhizal nodulin AGNOD-GHRP and
 demonstration of its toxicity to Escherichia coli."
 RT (in) Stacey G., Mullin B.C., Gresshoff P.M. (eds.);
 RL the Biology of Plant-Microbe Interactions:
 RL Proceedings of the 8th International Symposium on Molecular
 RL Plant-Microbe Interactions, pp.1-1, Unknown Publisher (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RA Twigg P.G.;
 RT "Isolation of a nodule-specific cDNA encoding a putative glycine-rich
 protein from Alnus glutinosa."
 RL Thesis (1993). The University of Tennessee, Knoxville, TN, USA.
 RN [3]
 RP SEQUENCE FROM N.A.

RC TISSUE-ROOT NODES;
 RA Pawlowski K., Twigg P.G., Dobritsa S.V., Guan C., Mullin B.C.;
 RT "A node-specific gene family from *Alnus glutinosa* encodes glycine
 RT and histidine-rich proteins expressed in the early stages of
 RT actinorhizal nodule development."
 RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U69156; AAD00171.1; -
 DR InterPro: IPR002395; -
 DR PRINTS: PR00334; KININOGEN.
 SQ SEQUENCE 99 AA; 10567 MW; 2ACBEAD57C070E83 CRC64;

Query Match 42.9%; Score 76; DB 10; Length 99;
 Best Local Similarity 48.1%; Pred. No. 0.0026;
 Matches 13; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

OY 1 GHKHHGHGHHKNGK--NGKNG 25
 ||:||||| | :||:
 DB 51 GHRHVHGHHGHHGNGNHGHHHG 77

RESULT 5
 O9LYB2 PRELIMINARY; PRT; 199 AA.
 AC O9LYB2;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHEICAL 21.5 KDA PROTEIN.
 GN T20010_200.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Obermayer B., Ottenwelder B., Duchemin D., Zeitler K., Mewes H.W.,
 RA Rudd S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL163816; CAB87755.1; -
 DR InterPro: IPR000345; -
 DR PRINTS: PR00334; KININOGEN.
 DR POSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR Hypothetical protein.
 KW SEQUENCE 199 AA; 21539 MW; E5D28AC167B3FBF8 CRC64;

Query Match 42.1%; Score 74.5; DB 10; Length 199;
 Best Local Similarity 46.4%; Pred. No. 0.0083;
 Matches 13; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

OY 1 GHKHHGHGHHKNGK--KNGKNGKNG 25
 ||:||||| | :||:
 DB 97 GHGHHGHHRRRRGRDRGRGHHHG 124

RESULT 6
 O39115 PRELIMINARY; PRT; 177 AA.
 AC O39115;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GLYCINE AND PROLINE RICH PROTEIN (GENOMIC DNA, CHROMOSOME 5, P1
 DE CLONE:MFC19).
 GN GPRP.
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. E. COLI LE392;
 RX MEDLINE=96189273; PubMed=8605310;
 RA Marty I., Monfort A., Stefel V., Ludevid D., Delseny M.,
 RA Puigdomenech P.;
 RT "Molecular characterization of the gene coding for GPRP, a class of
 RT proteins rich in glycine and proline interacting with membranes in
 RT Arabidopsis thaliana."
 RL Plant Mol. Biol. 30:625-636(1996).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. E. COLI LE392;
 RA Puigdomenech P.;
 RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=99397451; PubMed=10470850;
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
 RT PI and YAC clones."
 RL DNA Res. 6:183-195(1999).
 DR EMBL; X84315; CAA59059.1; -
 DR EMBL; AB018113; BAB09163.1; -
 DR Mendel: 17452; Arabid.2767;17452.
 SQ SEQUENCE 177 AA; 17830 MW; 3EF8094DCAD13F92 CRC64;

Query Match 41.0%; Score 72.5; DB 10; Length 177;
 Best Local Similarity 57.7%; Pred. No. 0.014;
 Matches 15; Conservative 3; Mismatches 5; Indels 3; Gaps 2;

OY 1 GHKHHGHGHHG--KHKNGK-KNGKH 23
 ||:||||| | :||:
 DB 140 GHGCGYGHGKFKGKFKGKH 165

RESULT 7
 O9LF59 PRELIMINARY; PRT; 173 AA.
 AC O9LF59;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GLYCINE/PROLINE-RICH PROTEIN.
 GN K10A8_130.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
 RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL391151; CAC01909.1; -
 DR InterPro: IPR000216; -
 DR PRINTS: PR00239; RHODOSPENTAL.
 SQ SEQUENCE 173 AA; 18536 MW; E3510947AA98BC0A CRC64;

Query Match 40.4%; Score 71.5; DB 10; Length 173;
 Best Local Similarity 39.1%; Pred. No. 0.018;
 Matches 18; Conservative 2; Mismatches 7; Indels 19; Gaps 3;

QY 1 GHGKHG-----GHGKHKKKKKK-----GKNG-----WK 27
 Db 128 GHGKHGCGYGYGHGKHGKFKRKGKRGKMGKFKRKKWK 173

RESULT 8
 ID 09W3L3 PRELIMINARY; PRT; 1937 AA.
 AC 09W3L3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE FEMALE STERILE (1) HOMOEOTIC PROTEIN.
 GN FS(1)H OR CG2252.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer J.R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Abmayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Bortone D.C., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL, AE003442; AAF6312.2;
 DR Flybase; FBgn0004656; fs(1)h.
 DR InterPro; IPR000104;
 DR InterPro; IPR001487;
 DR InterPro; IPR002173;
 DR InterPro; IPR002395;
 DR Pfam; PF00439; bromodomain; 2.
 DR PRINTS; PR00308; ANTIFREZEI.

DR PRINTS; PR00503; BROMODOMAIN.
 DR PRINTS; PR00334; KININOGEN.
 DR PROSITE; PS00633; BROMODOMAIN_1; 2.
 DR PROSITE; PS50014; BROMODOMAIN_2; 2.
 DR PROSITE; PS00583; PFRA_KINASES_L; UNKNOWN_1.
 SQ SEQUENCE 1937 AA; 19539 MW; 1D80AA7BB351F06B CRC64;

Query Match 40.1%; Score 71; DB 5; Length 1937;
 Best Local Similarity 46.2%; Pred. No. 0.23;
 Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 GHGKHGCGYGYGHGKHGKFKRKGKRGKMGKFKRKKWK 26
 Db 597 GHGKHGCGYGYGHGKHGKFKRKGKRGKMGKFKRKKWK 622

RESULT 9
 ID 09W4C1 PRELIMINARY; PRT; 554 AA.
 AC 09W4C1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CG15784 PROTEIN.
 GN CG15784.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer J.R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Abmayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Bortone D.C., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).

RT dssiation of the resurrection grass *Sporobolus stapfiannus*;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ242804; CAB61840.1; -
 DR InterPro: IPR000216; -
 DR PRINTS: PRO0239; RHODOPSNTAIL.
 SO SEQUENCE 197 AA; 19769 MW; 4EBD3097526A825A CRC64;

Query Match 39.0%; Score 69; DB 10; Length 197;
 Best Local Similarity 51.6%; Pred. No. 0.046;
 Matches 16; Conservative 1; Mismatches 10; Indels 4; Gaps 2;

OY 1 GH-KHKHGHC---HGKHKNGKNGKNGKNGK 27
 Db 167 GHGKFGHGHCGRHGKNGKNGKNGK 197

RESULT 14
 ID Q18577 PRELIMINARY; PRT; 251 AA.
 AC Q18577;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DE 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
 DE SIMILARITY TO THE RAT K-KININOGEN.
 GN C42D4.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smailon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Du Z., Le T.,
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.,
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U41991; AAA83352.1; -
 SO SEQUENCE 251 AA; 28352 MW; B066535FC510F419 CRC64;

Query Match 39.0%; Score 69; DB 5; Length 251;
 Best Local Similarity 44.0%; Pred. No. 0.058;
 Matches 11; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 GHKHKHGHCGRHGKNGKNGKNGKNGK 25
 Db 166 GPRHGHCGRHGKNGKNGKNGKNGK 190

RESULT 15
 ID Q9VU00 PRELIMINARY; PRT; 693 AA.
 AC Q9VU00;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
 DE CAUP PROTEIN.
 GN CAUP OR CG10605.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Fandel M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshireli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui Y., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RL "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER HOMEOBOX DOMAINS.
 DR EMBL: AE003540; AAF49895.1; -
 DR FlyBase; Fpgn0015919; caup.
 DR InterPro; IPR001356; -
 DR Pfam; PF00046; homeobox; 1.
 DR PROSITE; PS00027; HOMEOBOX_1; 1.
 DR PROSITE; PS00071; HOMEOBOX_2; 1.
 DR SMART; SM00389; HOX; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SO SEQUENCE 693 AA; 73667 MW; FBEB1616493F7EC9 CRC64;

Query Match 39.0%; Score 69; DB 5; Length 693;
 Best Local Similarity 48.0%; Pred. No. 0.16;
 Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

OY 1 GHKHKHGHCGRHGKNGKNGKNGKNGK 25
 Db 649 GHGSHGHCGRHGKNGKNGKNGKNGK 673

Search completed: July 6, 2001, 09:25:55
Job time: 991 sec

THIS PAGE BLANK (USPTO)